

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-185494

(43)Date of publication of application : 28.06.2002

(51)Int.Cl.

H04L 12/56

(21)Application number : 2000-374957

(71)Applicant : FUJITSU LTD

(22)Date of filing : 08.12.2000

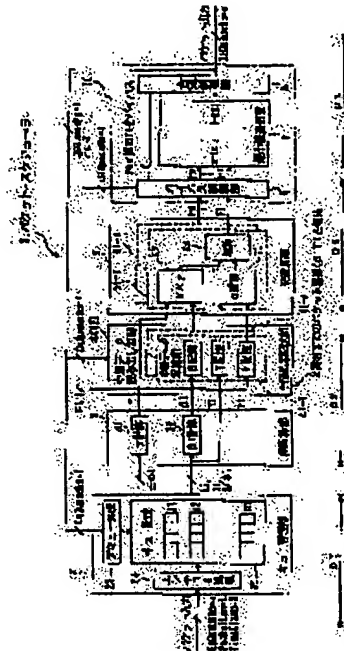
(72)Inventor : KYO EN

(54) PACKET SCHEDULER

(57)Abstract:

PROBLEM TO BE SOLVED: To increase largely the speed of a packet scheduling, by avoiding reduction of its throughput caused by the increase of the number of the stages of its pipe-line processing.

SOLUTION: At least first and second intermediate data are stored in an intermediate-calculation-result storing means 4. The first intermediate data are necessary for calculating the ideal outputting time of a first packet data to be outputted firstly from each queue. The second intermediate data are necessary for calculating the ideal outputting time of a second packet data to be so outputted that it follows the first packet data. Then, there is adopted, as the first intermediate data relative to a new first packet data of the queue, the second intermediate data relative to the second packet data belonging to the same queue as the first packet data which is determined by a transmission-packet-data determining means 10 that it must be outputted with the maximum priority. Subsequently, the second intermediate data are read out from the intermediate-calculation-result storing means 4, together with the first intermediate data relative to the first packet data belonging to other queues, as to output them thereafter to a calculating means 5.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2002-185494
(P2002-185494A)

(43) 公開日 平成14年6月28日 (2002.6.28)

(51) Int.Cl.⁷
H 0 4 L 12/56

識別記号

F I
H 0 4 L 11/20

テーマコード(参考)
1 0 2 C 5 K 0 3 0

審査請求 未請求 請求項の数 7 O L (全 28 頁)

(21) 出願番号 特願2000-374957(P2000-374957)

(22) 出願日 平成12年12月8日(2000.12.8)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 許 炎

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74) 代理人 100092978

弁理士 真田 有

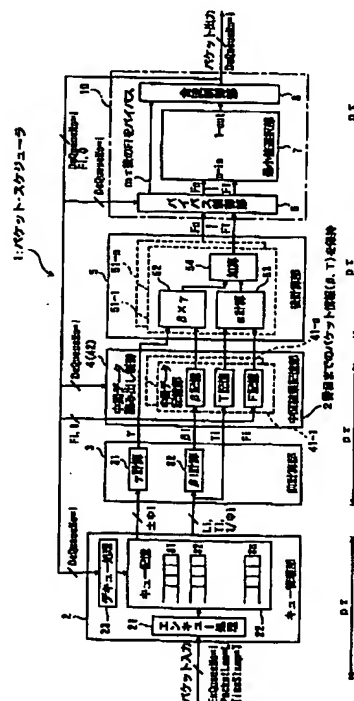
Fターム(参考) 5K030 HA08 KA03 LA03 LC01 LC09
LD18 LE03 LE05

(54) 【発明の名称】 パケット・スケジューラ

(57) 【要約】

【課題】 パイプライン処理の段数に起因して生じるスループットの低下を回避して、パケット・スケジューリングの大幅な高速化を図れるようにする。

【解決手段】 少なくとも、各キューから最初に出力されるべき第1パケットデータの理想出力時刻計算に必要な第1中間データと、その次に出力されるべき第2パケットデータについての理想出力時刻計算に必要な第2中間データとを中間計算結果記憶手段4に記憶しておき、送信パケットデータ決定手段10で最優先に送信すべきと決定された第1パケットデータと同じキューに属する第2パケットデータについての第2中間データを、当該キューの新たな第1パケットについての第1中間データとして他のキューに属する第1パケットの第1中間データとともに中間計算結果記憶手段4から読み出して後計算手段5へ出力するように構成する。



【特許請求の範囲】

【請求項1】 送信待ちのケットデータをそれぞれ出力保証帯域の設定が可能な複数のキューにより管理し該キュー毎のケットデータの送信量を該出力保証帯域の設定に基づいてスケジューリングするケット・スケジューラであって、

該ケットデータに関する情報に基づいて、空き状態のキューについての出力保証帯域を余剰帯域としてアクティブ状態の他のキューに公平に配分するための計算を行なう前計算手段と、

少なくとも、上記の各キューから最初に出力されるべき第1ケットデータの理想出力時刻計算に必要な該前計算手段での第1計算結果を含む第1中間データと、該第1ケットデータの次に出力されるべき第2ケットデータについての理想出力時刻計算に必要な該前計算手段での第2計算結果を含む第2中間データとを記憶する中間計算結果記憶手段と、

該中間計算結果記憶手段に記憶されている該第1中間データに基づいて上記の各キューにおける第1ケットデータについての理想送信時刻情報を計算する後計算手段と、

該後計算手段によって計算された各理想出力時刻情報に基づいて最優先に送信すべきキューの第1ケットデータを決定する送信ケットデータ決定手段と、

該送信ケットデータ決定手段で最優先に送信すべきと決定された第1ケットデータと同じキューに属する第2ケットデータについての第2中間データを、当該キューの新たな第1ケットデータについての第1中間データとして他のキューに属する第1ケットデータの第1中間データとともに該中間計算結果記憶手段から読み出して該後計算手段へ出力する中間計算結果読み出し制御手段とをそなえて構成されたことを特徴とする、ケット・スケジューラ。

【請求項2】 該中間計算結果読み出し制御手段が、該中間計算結果記憶手段に記憶されている、該送信ケットデータ決定手段で最優先に送信すべきと決定された第1ケットデータについての第1中間データを無効にするように構成されたことを特徴とする、請求項1記載のケット・スケジューラ。

【請求項3】 該中間計算結果記憶手段が、該第2ケットデータの次に出力されるべき第3ケットデータの理想出力時刻計算に必要な該前計算手段での第3計算結果を含む第3中間データを、新たな第2ケットについての第2中間データとして記憶するように構成されたことを特徴とする、請求項1又は請求項2に記載のケット・スケジューラ。

【請求項4】 該送信ケットデータ決定手段が、該後計算手段によって計算された理想出力時刻情報を比較して最優先に送信すべき第1ケットデータの理想出力時刻情報を選択する理想出力時刻情報選択手段と、

該理想出力時刻情報選択手段からの理想出力時刻情報と該後計算手段からバイパス入力される理想出力時刻情報とを比較して最優先に送信すべき第1ケットデータを決定する合流制御手段と、

該後計算手段において計算された該第2ケットデータの理想出力時刻情報を該バイパス入力として該合流制御手段へ出力するとともに、それ以外の第1ケットデータの理想出力時刻情報を該理想出力時刻情報選択手段へ出力するバイパス制御手段とをそなえて構成されていることを特徴とする、請求項1～3のいずれか1項に記載のケット・スケジューラ。

【請求項5】 該バイパス制御手段が、該キュー毎のバイパス制御回路をそなえるとともに、該バイパス制御回路が、それぞれ、該後計算手段によって計算された該第1ケットデータの理想出力時刻情報を記憶するメモリ部と、該メモリ部に記憶された該理想出力時刻情報を該バイパス入力もしくは該理想出力時刻情報選択手段への入力として選択的に出力するバイパス選択部と、

該合流制御手段で最優先に送信すべきと決定された第1ケットデータの属するキュー番号が自担当のキュー番号と一致すると該メモリ部の該理想出力時刻情報が該バイパス入力として選択され、該キュー番号が自担当のキュー番号と異なると該メモリ部の該理想出力時刻情報が該理想出力時刻情報選択手段への入力として選択されるよう該バイパス選択部を制御する状態制御部とをそなえて構成されたことを特徴とする、請求項4記載のケット・スケジューラ。

【請求項6】 該バイパス制御手段が、該合流制御手段で最優先に送信すべきと決定した第1ケットデータについての理想出力時刻情報を無効にするように構成されたことを特徴とする、請求項4記載のケット・スケジューラ。

【請求項7】 複数のケットインタフェースについての送信待ちのケットデータをそれぞれ出力保証帯域の設定が可能な複数のキューにより該ケットインタフェース毎に管理し該キュー毎のケットデータの送信量を該出力保証帯域の設定に基づいてスケジューリングするケット・スケジューラであって、

非アクティブ状態となっているキューについての出力保証帯域を余剰帯域としてアクティブ状態の他のキューに公平に配分するための計算を行なう前計算手段と、

少なくとも、上記の各キューから最初に出力されるべき第1ケットデータの理想出力時刻計算に必要な該前計算手段での第1計算結果を含む第1中間データと、該第1ケットデータの次に出力されるべき第2ケットデータの理想出力時刻計算に必要な該前計算手段での第2計算結果を含む第2中間データとを記憶する中間計算結果記憶手段と、

該中間計算結果記憶手段に記憶されている該第1中間デ

3

ータに基づいて上記の各キューにおける第1パケットデータの理想出力時刻情報を計算した後計算手段と、
該後計算手段によって計算された各理想出力時刻情報に基づいて最優先に送信すべきキューの第1パケットデータを決定する送信パケットデータ決定手段と、

該送信パケットデータ決定手段で最優先に送信すべきと決定された第1パケットデータと同じキューに属する第2パケットデータについての第2中間データを、当該キューの新たな第1パケットデータについての第1中間データとして他のキューに属する第1パケットデータにつ

いての第1中間データとともに該中間計算結果記憶手段から読み出して該後計算手段へ出力する中間計算結果読み出し制御手段とをそなえとともに、
上記の各手段を複数のパイプライン処理ステージに分割し、該パイプライン処理ステージが、それぞれ、一定時間毎に異なるパケットインタフェースについての処理を行なうように構成されたことを特徴とする、パケット・スケジューラ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、パケット・スケジューラに関し、特に、いわゆるルータなどのパケット転送装置においてWFQ(Weighted Fair Queuing)を用いたパケット・スケジューリングを行なうのに好適な、パケット・スケジューラに関する。

【0002】

【従来の技術】ネットワークにおけるパケットデータ(以下、単に「パケット」という)のフロー、あるいは、フローのアグリゲート(集合;以下、これらを総称してストリームという)のサービス品質(QoS:Quality of Service)を保証するために、IP(Internet Protocol)ルータやATM(Asynchronous Transfer Mode)ルータなどのネットワーク中継装置(パケット転送装置)では、それぞれのストリームに対して帯域を予約し、その予約帯域を保証するようにパケットの送信、即ち、各ストリーム毎のキューに保持されたパケットの読

$$F_{i,1} = \max\{F_{i,0}, T_{i,1}\} + \frac{L_{i,1}}{\Phi_i} \times \frac{\Phi_b}{R} \quad \dots (1-1)$$

【0007】

【数2】

$$Output = \min\{F_{i,1} \mid i \in active\} \quad \dots (1-2)$$

【0008】ここで、上記の式(1-1)、(1-2)で使用した各記号の意味を説明する。まず、式(1-1)の第1項において、Fはパケットの理想出力時刻、Tはパケットの到着時刻の情報をそれぞれ表わし、これらのFやTの添字部分(i,j)は、i番目のキュー(以下、キュー#iと表記する)の先頭からj番目のパケットを指す。即ち、F_{i,j}はキュー#iの先頭からj番目

4

み出し制御(パケット・スケジューリング)を行なう必要がある。なお、上記の各キューは、通常、先入れ先出し(FIFO:First-In First-Out)方式のメモリ(バッファ)で構成される。

【0003】一方、出力リンク(回線)の帯域を最大限に利用するため、各キューのアクティブ状況によって一時的に生じる剰余帯域をストリーム間で公平に配分する機能、即ち、非アクティブ状態のキュー(ストリーム)に対して設定されている予約帯域をアクティブ状態のキュー(ストリーム)に公平に配分する機能も必要となる。

【0004】なお、キューの「アクティブ状態」とは送信待ちのパケットがキューに存在している(保持されている)状態を意味し、「非アクティブ状態」とは、送信待ちのパケットがキューに存在しない(保持されていない、つまり空き)状態を意味する。以上のような機能をもつパケット・スケジューリング手法としては、WFQがよく知られているが、従来のWFQによる処理機能の実現では計算量が多いため、高速リンクに対応する実現が困難となっている。以下、WFQの計算原理について説明する。

【0005】パケット転送装置において、各ストリームのパケットはそれぞれ対応するキューに保持され、各キューに保持されたパケットはFIFO方式で出力される(先頭パケットが出力されると、次のパケットが先頭パケットとなる)ので、WFQでは、非アクティブ状態のキューの剰余帯域を公平に配分することも考慮した場合、各キューの先頭パケット(最初に出力されるべきパケット)の理想出力時刻F_{i,1}を次式(1-1)により計算し、得られた各キューの先頭パケットの理想出力時刻F_{i,1}に基づいて、最優先で出力すべき先頭パケットを次式(1-2)により選択して出力することが行なわれる。

【0006】

【数1】

の packets についての理想出力時刻(情報)を表わし、T_{i,j}はキュー#iの先頭からj番目のパケットの到着時刻(情報)を表わすことになる。

【0009】なお、1つのパケット・スケジューラ(以下、単に「スケジューラ」という)は1~n番目までのキューの処理を可能とするので、i=1~nである。また、j=1は、スケジューラで計算を行なっている時に、キュー#iの先頭に位置するパケットを表わし、j=0は直前に出力したパケットの出力時刻を表わす。つまり、式(1-1)の第1項max{F_{i,0}, T_{i,1}}は、直前に出力したキュー#iの先頭パケットの理想出

力時刻と、その先頭パケットに続く同じキュー# i のパケットの到着時刻とのうち値の大きい(遅い)方をとることを表わす。これは、先頭パケットの理想出力時刻 $F_{i,1}$ の計算基準となる時刻情報を、直前に出力したキュー# i の先頭パケットの出力時刻 $F_{i,0}$ と、その先頭パケットに続く同じキュー# i のパケットの到着時刻 $T_{i,1}$ とのいずれかに揃えることを意味する。

【0010】例えば、パケット出力後に次のパケットが到着する状況(次のパケット到着までにスケジューラによる計算が完了している状況)では、パケットの到着時刻 $T_{i,1}$ の方が直前に出力したパケットの出力時刻 $F_{i,0}$ よりも遅いので、次のパケットの到着時刻 $T_{i,1}$ が理想出力時刻 $F_{i,1}$ の計算基準となり、逆に、パケット出力前に次のパケットが到着する状況(スケジューラによる計算途中で次のパケットが到着する状況)では、直前に出力したパケットの出力時刻 $F_{i,0}$ の方が次のパケットの到着時刻 $T_{i,1}$ よりも遅いので、直前に出力したパケットの出力時刻 $F_{i,1}$ が理想出力時刻 $F_{i,1}$ の計算基準となる。

【0011】一方、式(1-1)の第2項において、 $L_{i,j}$ はキュー# i の先頭から j 番目のパケットの長さ、 Φi はキュー# i に設定された予約帯域、 Φb はアクティブ状態のキュー(以下、単に「アクティブキュー」ともいう) # i についての予約帯域 Φi の合計、 R は全キュー# $1 \sim n$ の予約帯域 Φi の合計をそれぞれ表わす。つまり、式(1-1)の第2項は、空き状態のキュー# i の予約帯域を他のキューへ公平に配分するためのパラメータを表わしていることになる。

【0012】そして、式(1-2)は、アクティブ状態のキューの中で最も理想出力時刻 $F_{i,1}$ の値が小さいものを選択することを意味する。つまり、理想出力時刻 $F_{i,1}$ が大きい先頭パケットほど、優先して出力すべき先頭パケットとしては選択されにくくなるのが分かる。例えば、或るキュー# i から1回の出力で大きなパケット長 L のパケットを出力した場合には、次のパケット(新たな先頭パケット)についての理想出力時刻 $F_{i,1}$ の値が大きくなるので、そのパケットは出力パケットとして選択されにくくなり、連続出力は行なわれにくくなる。

【0013】また、同じキュー# i から連続してパケットが出力された場合にも、式(1-1)の第2項の累積により、理想出力時刻 $F_{i,1}$ が増加してゆくので、或る程度連続してパケット出力が行なわれた時点で、そのキュー# i の先頭パケットは出力パケットとしては選択されにくくなる。逆に、例えば、予約帯域 Φi に大きな値が設定されているキュー# i は、他のキュー# i に比べて理想出力時刻 $F_{i,1}$ が小さくなるので、そのキュー# i のパケットが出力パケットとして選択されやすくなる。

【0014】なお、以下の説明において、便宜上、上記

の式(1-1)は次式(1-3)に示すように変形する。

$$F_{i,1} = \alpha_{i,1} + \beta_{i,1} \times \gamma \quad \dots (1-3)$$

ただし、 $\alpha_{i,1} = \max \{F_{i,0}, T_{i,1}\}$ 、 $\beta_{i,1} = L_{i,1} / \Phi i$ 、 $\gamma = \Phi b / R$ である。また、上記の $\alpha_{i,1}$ は「 α 」、 $\beta_{i,1}$ は「 β 」もしくは「 βi 」とそれぞれ略記する場合がある。同様に、前記の $F_{i,j}$ を「 F 」もしくは「 $F i$ 」、 $T_{i,j}$ を「 T 」もしくは「 $T i$ 」とそれぞれ略記する場合もある。以上の記号は、後述する本発明の実施形態においても同様の意味で使用するものとする。

【0015】さて、ここで、上述したような計算を実際にハードウェア(パケット・スケジューラ)として実現するには、例えば、図21に示すような構成が考えられる。即ち、この図21に示すように、WFQを用いたパケット・スケジューラ100は、キュー管理部101、計算部102及び最小値選択部103をそなえて構成され、さらに、キュー管理部101は、エンキュー処理部111、キュー記憶部112及びデキュー処理部113をそなえ、計算部102は、 γ 計算部121、 β 計算部122、各キュー# i 毎の記憶部123-1~123-n及び各キュー毎の F 計算部124-1~124-nをそなえて構成されている。

【0016】ここで、キュー管理部101において、キュー記憶部112は、図示しないパケットメモリに格納された到着パケットに関する情報(パケット長 L 、到着時刻 T (TimeStamp)、上記パケットメモリでの格納位置(メモリアドレス; TopPointer))を、その到着パケットが属するストリーム対応のキュー# $1 \sim n$ 毎に $F I F O$ 形式で記憶してキュー管理テーブルとして管理するためのものである。

【0017】また、エンキュー処理部111は、上記のパケットメモリから通知されるエンキュー番号(EnQueueNo)に基づいて、上記の到着パケットの情報を、キュー記憶部112(キュー管理テーブル)の対応するキュー# i に格納(エンキュー)するものであり、デキュー処理部113は、最小値選択部103にて最優先に出力すべきと決定した先頭パケットの属するキュー番号 i (デキュー指示; DeQueueNo= i) を受けることにより、

そのキュー番号 i に対応するキュー# i の先頭パケットの情報(メモリアドレス)をパケットメモリに通知して該当パケットを出力リンク(回線)に送出させるとともに、そのパケットの情報(パケット長 L 、到着時刻 T (TimeStamp)、メモリアドレス (TopPointer))を削除するものである。

【0018】なお、本キュー管理部101では、各キュー# i 毎に設定される予約帯域 Φi やアクティブキュー# i の合計予約帯域 Φb 、最小値選択部103で選択された先頭パケットの出力時刻 $F_{i,0}$ なども管理されている。次に、計算部102において、 γ 計算部121は、

キュー管理部101で管理されているアクティブキュー# i の合計予約帯域 Φb を基に前記の γ ($=\Phi b/R$)を計算するためのものであり、 β 計算部122は、キュー管理部101で管理されている情報のうち、バケット長 L と予約帯域 Φi とにより上記の β ($=L/\Phi i$)を計算するためのもので、その計算結果 β は、キュー番号 i に対応する記憶部123- i に記憶されるようになっている。なお、キュー管理部101で管理されている到着時刻 T 及び出力時刻 $F_{i,0}$ は、 β 計算部122は経由せずにそのまま対応するキュー番号 i の記憶部123- i にて上記の計算結果 β と組で記憶される。

【0019】また、各F計算部124- i は、それぞれ、上記の記憶部123に記憶されている情報(β , T , $F_{i,0}$)に基づいて、前記の式(1-3)による計算を行なうことにより、対応するキュー# i の先頭バケットについての理想出力時刻 F を計算する $F_{i,1}$ を計算するためのもので、このために、図21中に示すように、それぞれ、 $\beta \times \gamma$ 計算部125、 α 計算部126及びこれらの各計算部125、126の計算結果を加算することにより理想出力時刻 $F_{i,1}$ を求める加算部127を有して構成されている。

【0020】そして、上記の最小値選択部103は、上記の各F計算部124- i によって求められた各キュー# i 毎の先頭バケットの理想出力時刻 $F_{i,1}$ を、例えば、トーナメント方式などで比較することにより、最小の理想出力時刻 $F_{i,1}$ をもつ先頭バケットの属するキュー# i を選択(決定)するためのもので、そのキュー番号 i がデキュー指示としてデキュー処理部113に通知(フィードバック)されることによって、該当するキュー# i の先頭バケットがバケットメモリから読み出されて出力リンクへ転送されるようになっている。なお、このとき選択された理想出力時刻 $F_{i,1}$ も、キュー管理部101に通知され、キュー管理部101にて、次の計算で使用する出力時刻 $F_{i,0}$ として管理されることになる。

【0021】上述のごとく構成されたバケット・スケジューラ100では、或るストリームからのバケット到着毎に、エンキュー処理部111によって、その到着バケットが属するストリームに対応するキュー# i の状態が更新されてゆく。そして、キュー# i の先頭にバケットが存在する場合、その先頭バケットに関する情報(バケット長 $L_{i,1}$, 到着時刻 $T_{i,1}$), 直前に出力された先頭バケットに関する情報(出力時刻 $F_{i,0}$)及びキュー# i 対応の設定情報($1/\Phi i$)がそれぞれキュー管理部101から計算部124に送られる。

【0022】計算部124- i では、まず、上述のごとくキュー管理部101から送られてきた情報により、 γ 計算部121にて γ の値が計算されるとともに、 β 計算部122にて β の値が計算され、 β は対応するキュー番号 i の記憶部123- i に保持される。なお、 β の計算

に用いられなかった情報(到着時刻 $T_{i,1}$, 出力時刻 $F_{i,0}$)は、それぞれ、 β 計算部122によって得られた β とともに記憶部123- i に保持される。

【0023】そして、 γ 計算部121によって求められた γ と、記憶部123- i に保持された β , $T_{i,1}$ 及び $F_{i,0}$ とは、それぞれ、同じタイミングで対応するキュー番号 i のF計算部124- i に入力され、そのF計算部124- i では、このようにして送られてきた情報(γ , β , $T_{i,1}$ 及び $F_{i,0}$)を基に、 $\beta \times \gamma$ 計算部125にて $\beta \times \gamma$ の値が計算されるとともに、 α 計算部126にて α の値が計算され、それらの計算結果が加算部127にて加算されることにより、キュー番号 i の先頭バケットについての理想出力時刻 $F_{i,1}$ が求められる。

【0024】このようにして各F計算部124- i にて求められた各理想出力時刻 $F_{i,1}$ は、それぞれ、同一タイミングで最小値選択部103に入力され、最小値選択部103は、トーナメント方式などによって各理想出力時刻 $F_{i,1}$ を比較して、最も値の小さい理想出力時刻 $F_{i,1}$ を選択し、そのキュー番号 i をデキュー指示としてキュー管理部101のデキュー処理部113に通知する。

【0025】これにより、デキュー処理部113によって、デキュー番号 i のキュー# i の先頭バケットが読み出されて、出力リンクに送出される。以降、同様に、各キュー# i 毎に先頭バケットの理想出力時刻 $F_{i,1}$ が計算されて、最も小さい理想出力時刻 $F_{i,1}$ をもつキュー# i の先頭バケットが最優先に出力すべきバケットとして選択されて、出力リンクへ送出されてゆく。

【0026】ただし、本バケット・スケジューラ100は、図21中に示すように、キュー管理部101、記憶部123- i (γ 計算部121), F計算部124- i 及び最小値選択部103が、それぞれ、時間とともに同時に動作し一定時間(一定クロック数) τ 毎に情報を後段へ出力する4段のパイプライン構成になっており、計算部102では、次のデキュー指示が発生するまで、同じ情報に基づいた計算が繰り返行なわれ、次のデキュー指示が発生すると、パイプライン処理によるそれまでの途中計算結果を破棄する(無効にする)ことで、デキュー指示発生後の新たな情報に基づく計算が開始されるようになっている。

【0027】また、上述した計算過程において、バケットの入出力によってキュー# i のアクティブ状態に変化があると、合計予約帯域 Φb の値が変化する(つまり、アクティブキュー# i に割り当て可能な余剰帯域の値が変化する)ため、式(1-3)の再計算が必要となる。即ち、空き状態のキュー# i にバケットが新たに格納されてアクティブキュー# i の数が増えたり、1つしかバケットが保持されていないキュー# i からそのバケットが出力されて空き状態のキュー# i が増えたりすると、合計予約帯域 Φb の値が変化するため、 γ 計算部121

において、 $\gamma (= \Phi b / R)$ の再計算が必要になるのである。

【0028】このため、上記の γ 計算部121は、空き状態からアクティブ状態に移したキュー#iが生じると上記の合計予約帯域 Φb として $\Phi b = \Phi b + \Phi i$ を用いて γ の再計算を行ない、逆に、アクティブ状態から空き状態に移したキュー#iが生じると上記の合計予約帯域 Φb として $\Phi b = \Phi b - \Phi i$ を用いて γ の再計算を行なう。

【0029】以上のようにして、バケット・スケジューラ100は、各キュー#iのアクティブ状況によって一時的に生じる剰余帯域をストリーム間で公平に配分しながら、設定された予約帯域 Φi を保証するようにバケットの送信(バケット・スケジューリング)を行なうことが可能である。

【0030】

【発明が解決しようとする課題】しかしながら、このような従来のバケット・スケジューラ100では、デキュー指示の発生からの一連のパイプライン処理の合計遅延時間が、そのスループットを決定してしまう。例えば、パイプライン段数をm段(図21では $m=4$ 段)とすると、1パイプライン当たりの処理時間(遅延時間)は上記のように $p\tau$ なので、上記の合計遅延時間は $m \times p\tau$ と表わせる。

【0031】ここで、最小バケット長(L_{min})のバケットが連続して入出力するときのスループット(S)は、 $S = L_{min} / (m \times p\tau)$ によって求められる。つまり、バケット・スケジューラ100のスループット(S)は、パイプライン処理の総段数(m)によって制限されてしまうのである。そこで、例えば、 $m = L_{min} / (S \times p\tau)$ の時間間隔で各キュー#iの先頭バケットに関する情報をパイプライン処理に入力することで、スループット(S)の向上を図ることも考えられるが、パイプライン処理では、上述したように、同じバケットについてのデキュー指示が重複して発生しないよう、必ず、デキュー指示によってパイプライン処理中のキュー#iの先頭バケットに関する情報を無効にしなければならないので、実際には、同じキュー#iからは連続的に(つまり、 $p\tau$ 間隔で)バケットを出力(デキュー)することができない。

【0032】即ち、或るキュー#iの先頭バケットに対するデキュー指示が発生した場合、他のキュー#iの先頭バケット(最小値選択部103で選択されなかったキュー#iの先頭バケット)のパイプライン処理による計算途中の情報については有効であるが、デキュー済み先頭バケットについてのパイプライン処理による計算途中の情報については無効にしなければならない。

【0033】このため、次に同じキュー#iからバケットがデキューされるのは、キュー管理部101から次の先頭バケットについての情報が計算部102に新たに送

られて、新たに理想出力時刻 $F_{i,1}$ が求められ、その値が最小値選択部103にて比較・選択されるまでなく、結果として、同じキュー#iからは $(m-1) \times p\tau$ 間隔でしかバケットをデキューすることができず、バケット出力のスループットが低下してしまうのである。

【0034】本発明は、このような課題に鑑み創案されたもので、パイプライン処理の段数に起因して生じるスループットの低下を回避して、バケット・スケジューリングの大幅な高速化を図れるようにした、バケット・スケジューラを提供することを目的とする。

【0035】

【課題を解決するための手段】上記の目的を達成するために、本発明のバケット・スケジューラ(請求項1)は、送信待ちのバケットデータをそれぞれ出力保証帯域の設定が可能な複数のキューにより管理し各キュー毎のバケットデータの送信量をそれぞれの出力保証帯域の設定に基づいてスケジューリングするものであって、次の各手段をそなえて構成されていることを特徴としている。

【0036】(1) 上記のバケットデータに関する情報に基づいて、空き状態のキューについての出力保証帯域を剰余帯域としてアクティブ状態の他のキューに公平に配分するための計算を行なう前計算手段

(2) 少なくとも、上記の各キューから最初に出力されるべき第1バケットデータの理想出力時刻計算に必要な上記前計算手段での第1計算結果を含む第1中間データと、上記第1バケットデータの次に出力されるべき第2バケットデータについての理想出力時刻計算に必要な上記前計算手段での第2計算結果を含む第2中間データとを記憶する中間計算結果記憶手段

(3) この中間計算結果記憶手段に記憶されている第1中間データに基づいて上記の各キューにおける第1バケットデータについての理想送信時刻情報を計算する後計算手段

(4) この後計算手段によって計算された各理想出力時刻情報に基づいて最優先に送信すべきキューの第1バケットデータを決定する送信バケットデータ決定手段

(5) この送信バケットデータ決定手段で最優先に送信すべきと決定された第1バケットデータと同じキューに属する第2バケットデータについての第2中間データを、そのキューの新たな第1バケットデータについての第1中間データとして他のキューに属する第1バケットデータの第1中間データとともに上記の中間計算結果記憶手段から読み出して後計算手段へ出力する中間計算結果読み出し制御手段

ここで、上記の中間計算結果読み出し制御手段は、上記の中間計算結果記憶手段に記憶されている、上記送信バケットデータ決定手段で最優先に送信すべきと決定された第1バケットデータについての第1計算結果を無効にするように構成されるのが好ましい(請求項2)。

【0037】また、上記の中間計算結果記憶手段は、上記の第2パケットデータの次に出力されるべき第3パケットデータの理想出力時刻計算に必要な上記前計算手段での第3計算結果を含む第3中間データを、新たな第2パケットについての第2中間データとして記憶するように構成されていてもよい（請求項3）。さらに、上記の送信パケットデータ決定手段は、次の各手段をそなえて構成されていてもよい（請求項4）。

【0038】（1）上記の後計算手段によって計算された理想出力時刻情報を比較して最優先に送信すべき第1パケットデータの理想出力時刻情報を選択する理想出力時刻情報選択手段

（2）この理想出力時刻情報選択手段からの理想出力時刻情報と上記の後計算手段からバイパス入力される理想出力時刻情報とを比較して最優先に送信すべき第1パケットデータを決定する合流制御手段

（3）上記の後計算手段において計算された第2パケットデータの理想出力時刻情報を上記のバイパス入力として上記の合流制御手段へ出力するとともに、それ以外の第1パケットデータの理想出力時刻情報を上記の理想出力時刻情報選択手段へ出力するバイパス制御手段

なお、上記のバイパス制御手段は、上記キュー毎のバイパス制御回路をそなえるとともに、これらの各バイパス制御回路が、それぞれ、次の各部をそなえていてもよい（請求項5）。

【0039】（1）上記の後計算手段によって計算された第1パケットデータの理想出力時刻情報を記憶するメモリ部

（2）このメモリ部に記憶された理想出力時刻情報を上記バイパス入力もしくは上記理想出力時刻情報選択手段への入力として選択的に出力するバイパス選択部

（3）上記の合流制御手段で最優先に送信すべきと決定された第1パケットデータの属するキュー番号が自担当のキュー番号と一致すると上記メモリ部の理想出力時刻情報が該バイパス入力として選択され、上記キュー番号が自担当のキュー番号と異なると上記メモリ部の理想出力時刻情報が上記理想出力時刻情報選択手段への入力として選択されるよう上記バイパス選択部を制御する状態制御部

また、上記のバイパス制御手段は、上記の合流制御手段で最優先に送信すべきと決定した第1パケットデータについての理想出力時刻情報を無効にするように構成されるのが好ましい（請求項6）。

【0040】さらに、本発明のパケット・スケジューラは、複数のパケットインタフェースについての送信待ちのパケットデータをそれぞれ出力保証帯域の設定が可能な複数のキューにより上記の各パケットインタフェース毎に管理し該キュー毎のパケットデータの送信量を上記出力保証帯域の設定に基づいてスケジューリングするパケット・スケジューラであって、次の各手段をそなえる

とともに、これらの各手段が複数のパイプライン処理ステージに分割され、且つ、これらのパイプライン処理ステージが、それぞれ、一定時間毎に異なる上記パケットインタフェースについての処理を行なうように構成されたことを特徴としている（請求項7）。

【0041】（1）非アクティブ状態となっているキューについての出力保証帯域を余剰帯域としてアクティブ状態の他のキューに公平に配分するための計算を行なう前計算手段

（2）少なくとも、上記の各キューから最初に出力されるべき第1パケットデータの理想出力時刻計算に必要な該前計算手段での第1計算結果を含む第1中間データと、上記第1パケットデータの次に出力されるべき第2パケットデータの理想出力時刻計算に必要な該前計算手段での第2計算結果を含む第2中間データとを記憶する中間計算結果記憶手段

（3）この中間計算結果記憶手段に記憶されている上記第1中間データに基づいて上記の各キューにおける第1パケットデータの理想出力時刻情報を計算する後計算手段

（4）この後計算手段によって計算された各理想出力時刻情報に基づいて最優先に送信すべきキューの第1パケットデータを決定する送信パケットデータ決定手段

（5）この送信パケットデータ決定手段で最優先に送信すべきと決定された第1パケットデータと同じキューに属する第2パケットデータについての第2中間データを、当該キューの新たな第1パケットデータについての第1中間データとして他のキューに属する第1パケットデータについての第1中間データとともに上記の中間計算結果記憶手段から読み出して後計算手段へ出力する中間計算結果読み出し制御手段

【0042】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態を説明する。

（A）一実施形態の説明

図1は本発明の一実施形態としてのパケット・スケジューラの構成を示すブロック図で、この図1に示すパケット・スケジューラ1は、キュー管理部2、前計算部3、中間結果記憶部4、後計算部5及び最小値決定処理部10（バイパス制御部6、最小値選択部7及び合流制御部8）をそなえて構成されている。なお、本パケット・スケジューラ1は、この図1中に示すように、それぞれ p と τ 間隔で後段の処理へ情報を出力する4段のパイプライン構成になっているものとする。

【0043】ここで、キュー管理部2は、図21により前述したものと同様に、送信待ちのパケットデータ（以下、単に「パケット」という）をそれぞれ予約帯域（出力保証帯域） Φi の設定が可能な複数のキュー $\#1 \sim \#n$ により管理するためのもので、このために、本実施形態では、図1及び図2に示すように、エンキュー処理部

21、キュー記憶部22、デキュー処理部23及びエンキュー・デキュー調停部24をそなえて構成されている。なお、図1では、エンキュー・デキュー調停部24の図示は省略している。

【0044】そして、上記のキュー記憶部22は、本実施形態においても、到着パケットに関する情報〔関連パラメータ；パケット長L、到着時刻T(TimeStamp)、図2に示すパケットメモリ9での格納位置(メモリアドレス；TopPointer)〕を、その到着パケットが属するストリーム対応のキュー#1～#n毎にFIFO形式で記憶してキュー管理テーブル221(図2参照)として管理するためのものである。

【0045】また、エンキュー処理部21は、パケットメモリ9から通知されるエンキュー番号(EnQueueNo)に基づいて、上記の到着パケットの情報(関連パラメータ)を、キュー管理テーブル221の対応するキュー#iに格納(エンキュー)することにより、パケット到着(パケットメモリ9への書き込み)毎にその状態を更新するものである。

【0046】ただし、本エンキュー処理部21は、最初の計算では各キュー#iの先頭から1番目のパケットと2番目のパケットの情報をキュー管理テーブル221から読み出してそれぞれ時系列に後段の前計算部3に出力するようになっている。また、上記のパケット情報のうち到着時刻T(TimeStamp)については、図2では、エンキュー処理部21がキュー管理テーブル221に対してエンキューを行なったときのタイマ25でのタイマ値として得られる構成になっている。

【0047】一方、デキュー処理部23は、後述するデキュー指示(デキュー番号i)を受けることにより、そのデキュー番号iに対応する先頭パケットの情報(メモリアドレス；TopPointer)をキュー管理テーブル221から読み出してパケットメモリ9に通知するもので、これにより、そのメモリアドレスに格納されているパケットがパケットメモリ9から読み出されて出力リンク(回線)に送出されることになる。

【0048】ただし、本実施形態のデキュー処理部23は、上記のデキュー指示を受けると、デキュー番号iに対応するキュー管理テーブル221に格納されているデキュー済みのパケット情報を削除するとともに、同じキュー#iの3番目のパケット情報をキュー管理テーブル221から読み出して前計算部3に出力する機能も有している。

【0049】また、エンキュー・デキュー調停部24は、少なくとも、キュー管理テーブル221の同じキュー#iに同一タイミングで上記のエンキュー処理部21及びデキュー処理部23がアクセスしてパケットの情報が消失してしまうようなことを防止するために、エンキュー処理部21及びデキュー処理部23のキュー管理テーブル221に対するアクセスタイミングをそれぞれ調

停するためのものである。

【0050】次に、前計算部3は、上記のキュー管理部101から出力されてくるパケットの情報に基づいて、WFQの計算、即ち、空き状態のキュー#iについての予約帯域 Φ_i を余剰帯域としてアクティブキュー#iに公平に配分するための計算を行なうためのもので、このために、図1に示すように、 γ 計算部31及び β 計算部32をそなえて構成されている。

【0051】ここで、 γ 計算部31は、余剰帯域の配分変数 $[\gamma = \Phi_b \times (1/R)]$ を計算するためのものである。ただし、パケットの入出力でキュー#iのアクティブ状況が変化した場合には、 γ 計算部31は、 γ の再計算を行なう。例えば、 Φ_b がアクティブキューの予約帯域 Φ_i を合計した値で、初期値(全てのキュー#iにパケット情報が存在しないとき)が0であるとする、空き状態からアクティブ状態に移移したキュー#iが生じると、 $\Phi_b = \Phi_b + \Phi_i$ を計算しその Φ_b で γ を計算し直す。一方、アクティブ状態から空き状態に移移したキュー#iが生じると、 $\Phi_b = \Phi_b - \Phi_i$ を計算しその Φ_b で γ を計算し直す。

【0052】なお、本 γ 計算部31は、例えば図2に示すように、上記の Φ_i の代わりに、 (Φ_i/R) を予めRAM311に設定しておけば、その値 (Φ_i/R) をアクティブキュー#j($j \in \text{active}$)の数だけ γ 演算器(加算器)312にて累積加算してゆくだけで、 $\Phi_b \times (1/R)$ を求めることができ、乗算の計算を省略することができる。

【0053】一方、上記の β 計算部32は、パケット長(L)×予約帯域 $(1/\Phi_i)$ の計算により前記の β を求めるもので、例えば図2では、エンキュー処理部21から通知されるキュー番号(QueueNo=i)により、予めRAM321に設定された $(1/\Phi_i)$ を読み出して、その値 $(1/\Phi_i)$ と、同じくエンキュー処理部21から通知されるパケット長L(PacketLen)とを乗算器322にて乗算することにより、 β を求める構成になっている。

【0054】次に、中間結果記憶部4は、各キュー#iの先頭と2番目のパケットの情報に対しての中間計算データ $(\beta_{i,1}, T_{i,1})$ 、 $(\beta_{i,2}, T_{i,2})$ と、合流制御部8の出力からフィードバックされてくる、直前に出力した先頭パケットの出力時刻 $(F_{i,0})$ とをキュー#i毎に保持して後段の後計算部5へ出力するためのものである。なお、以下では、これらの中間計算データ $(\beta_{i,1}, T_{i,1})$ 、 $(\beta_{i,2}, T_{i,2})$ と、直前に出力した先頭パケットの理想出力時刻 $(F_{i,0})$ とを総称して「中間データ」という場合がある。

【0055】つまり、上記の中間結果記憶部4は、少なくとも、上記の各キュー#iの先頭パケット(第1パケットデータ)の理想出力時刻計算Fに必要な前計算部31での第1計算結果として上記の $\beta_{i,1}$ を含む(第1)

中間データと、上記先頭バケットの次に出力されるべき 2 番目のバケット (第 2 バケットデータ) についての理想出力時刻計算に必要な第 2 計算結果として上記の $\beta_{i,2}$ を含む (第 2) 中間データとを記憶するようになっているのである。

【0056】このため、本中間結果記憶部 4 は、図 1 及び図 2 に示すように、それぞれ、上記の中間データを記憶するためのキュー毎 # i の中間データ記憶部 41-1 ~ 41- n と、デキュー指示により上記中間データの読み出し制御を行なう中間データ読み出し制御部 42 とをそなえて構成されている。なお、図 2 では、中間データ読み出し制御部 42 の図示は省略している。

【0057】具体的に、この中間データ読み出し制御部 42 は、デキュー指示 (DeQueueNo= i) を受けることにより、そのデキュー番号 i と対応する中間データ記憶部 41-1 の 2 番目のバケット情報 ($\beta_{i,2}, T_{i,2}$) と出力時刻 ($F_{i,0}$) とを読み出して後計算部 5 へ出力するとともに、中間データ記憶部 41- i において、その 2 番目バケットの情報を先頭バケットの情報として更新し、且つ、合流制御部 8 の出力からデキュー指示 (DeQueueNo= i) とともにフィードバックされてくる出力時刻

($F_{i,0}$) を記憶する一方、デキュー指示 (DeQueueNo= i) 以外のキュー # k ($k \neq i$) については、先頭のバケット情報 ($\beta_{k,1}, T_{k,1}$) を対応する中間データ記憶部 41- k から読み出して後計算部 5 に出力するようになっている。

【0058】つまり、本中間データ読み出し制御部 42 は、後述する合流制御部 8 で最優先に送信 (デキュー) すべきと決定された先頭バケットと同じキュー # i に属する 2 番目のバケットについての中間データを、そのキュー # i の新たな先頭バケットについての中間データとして他のキュー # k に属する先頭バケットの中間データとともに中間計算結果記憶部 4 から読み出して後計算部 5 へ出力するようになっているのである。これにより、一定の遅延時間の後段処理のためのデータ出力が可能となり、デキューから次のバケットの計算開始までの遅延時間を短縮することが可能である。

【0059】なお、上述のごとく 2 番目バケットの情報を新たな先頭バケットの情報として更新したキュー # i については、キュー管理部 2 から送出された 3 番目のバケットの情報に基づいて前計算部 3 で計算されたバケット情報 ($\beta_{i,2}, T_{i,2}$) と、デキュー指示 (DeQueueNo= i) とともに送られてくる直前に出力した先頭バケットについての出力時刻 ($F_{i,0}$) とが新たな 2 番目のバケットについての中間データとして中間データ記憶部 41- i に記憶される。

【0060】つまり、本バケット・スケジューラ 1 は、一定時間間隔 $p \tau$ 毎に、2 番目バケットの情報を新たな先頭バケットの情報として更新したことで空きのできた中間データ記憶部 41- i に、その次 (3 番目) のバケ

ットについての中間データを、順次、補填してゆくことで、常に、先頭バケットと 2 番目のバケットとの 2 つ分のバケットについての中間データを中間データ記憶部 41- i に保持しておくように動作するのである。

【0061】次に、後計算部 5 は、パイプライン前段の上記の中間結果記憶部 4 から入力された、全アクティブキュー # i の中間データに基づいて、各アクティブキュー # i 毎に、前記の式 (1-3) による演算を行なうて、各アクティブキュー # i の先頭バケットの理想出力時刻 $F_{i,1}$ を求めるためのもので、このために、図 1 及び図 2 に示すように、キュー # i 毎の F 計算部 51-1 ~ 51- n を有しており、これらの各 F 計算部 51- i に、それぞれ、 $\beta \times \gamma$ を計算する $\beta \times \gamma$ 計算部 (乗算器) 52 と、 $\alpha_{i,1} = \max \{F_{i,0}, T_{i,1}\}$ を計算する α 計算部 53 と、これらの各計算部 52、53 の計算結果を加算することにより理想出力時刻 $F_{i,1}$ を求める加算部 54 とがそなえられている。

【0062】また、最小値決定処理部 (送信バケットデータ決定手段) 10 は、上記の後計算部 5 によって計算された各理想出力時刻 $F_{i,1}$ に基づいて最優先に送信すべきキュー # i の先頭バケットを決定 (選択) するためのもので、このために、本実施形態では、それぞれ、次のような機能を有する前記バイパス制御部 6、最小値選択部 7 及び合流制御部 8 がそなえられている。

【0063】即ち、バイパス制御部 6 は、各キュー # i と対応して、後計算部 5 からの計算結果 (キュー # i 毎の理想出力時刻 $F_{i,1}$) を最小値選択部 7 を経由せずに合流制御部 8 へバイパスするかどうかを示す状態変数 S_i を有し、この状態変数 S_i に応じて理想出力時刻 $F_{i,1}$ を最小値選択部 7 及び合流制御部 8 のいずれかに選択的に出力するためのもので、例えば、 $S_i = ON$ 状態の場合には、次に後計算部 5 から入力されてくる理想出力時刻 $F_{i,1}$ を合流制御部 8 へバイパスし、 $S_i = OF$ 状態の場合には、バイパスせずに最小値選択部 7 へ出力するようになっている。

【0064】このため、本バイパス制御部 6 には、図 2 に示すように、各キュー # i 毎のバイパス制御回路 61-1 ~ 61- n がそなえられており、さらに、これらの各バイパス制御回路 61- i には、それぞれ、例えば図 3 に示すように、一時記憶部 62、選択部 63 及び状態制御部 64 が設けられている。ここで、上記の一時記憶部 (メモリ部) 62 は、後計算部 5 の対応する F 計算部 51- i から入力される理想出力時刻 $F_{i,1}$ を一時的に記憶しておくためのものであり、選択部 (バイパス選択手段) 63 は、この一時記憶部 62 に記憶された理想出力時刻 $F_{i,1}$ を合流制御部 8 へのバイパス入力もしくは最小値選択部 7 への入力として選択的に出力するためのものである。

【0065】そして、状態制御部 64 は、例えば図 4 に示す状態遷移図に従って状態変数 S_i を制御すること

で、選択部63での選択状態を制御するためのものである。即ち、初期状態において状態変数 S_i はOFF状態（非バイパス状態；状態遷移65-1）であり、OFF状態で、自担当のキュー番号 i と一致するデキュー番号 i を受けると、状態変数 S_i はON状態（バイパス状態）に遷移して（状態遷移65-2）、状態制御部64は、一時記憶部62への上書きを許容することで出力待機中の先頭パケットについての理想出力時刻 $F_{i,1}$ を破棄する（無効にする）とともに、次に後計算部5から到着する先頭パケット（元々は2番目のパケット）についての理想出力時刻 $F_{i,1}$ がバイパスされる状態に選択部63を制御する。

【0066】一方、状態変数 S_i =OFF状態のときに受けたデキュー番号 i が自担当のキュー番号 i と一致しない場合は、状態変数 S_i はOFF状態のまま（状態遷移65-3）で、状態制御部64は、一時記憶部62において待機中の先頭パケットについての理想出力時刻 $F_{i,1}$ が最小値選択部7へ出力されるよう選択部63を制御する。なお、状態変数 S_i =OFF状態で、理想出力時刻 $F_{i,1}$ が単に入力されてきた場合も、状態変数 S_i は変化せず（OFF状態のまま）、入力されてきた理想出力時刻 $F_{i,1}$ が一時記憶部62に記憶される（状態遷移65-4）。

【0067】また、状態変数 S_i =ON状態のときに受けたデキュー番号 i が自担当のキュー番号 i と一致した場合は、状態変数 S_i は変化せず（ON状態のまま）、状態制御部64は、一時記憶部62において待機中の先頭パケットについての理想出力時刻 $F_{i,1}$ がバイパスされるよう選択部63を制御する（状態遷移65-5）。

【0068】一方、状態変数 S_i =ON状態のときに受けたデキュー番号 i が自担当のキュー番号 i と一致しない場合は、状態変数 S_i はOFF状態に遷移し、状態制御部64は、一時記憶部62において待機中の先頭パケットについての理想出力時刻 $F_{i,1}$ が最小値選択部7へ出力されるよう選択部63を制御する（状態遷移65-6）。

【0069】また、状態変数 S_i =ON状態のときに、理想出力時刻 $F_{i,1}$ が入力されてきた場合は、状態変数 S_i は変化せず（OFF状態のまま）、その理想出力時刻 $F_{i,1}$ がそのまま一時記憶部62に記憶されるとともに、状態制御部64が、その理想出力時刻 $F_{i,1}$ を合流制御部8へバイパスするよう選択部63を制御する（状態遷移65-7）。

【0070】なお、キュー# i が空き状態となり、処理すべきデータ（理想出力時刻 $F_{i,1}$ ）が存在しなくなった場合は、状態変数 S_i （以下、単に「 S 」と略記する場合もある）は、元の状態（ON/OFF状態）に関わらず、OFF状態となり、状態制御部64は、最小値選択部7への出力が選択されるよう選択部63を制御する（状態遷移65-8、65-9）。

【0071】このようにして、本状態制御部64は、自担当のキュー番号 i と一致するデキュー番号 i のデキュー指示を受けると、次に後計算部5（F計算部51-i）から到着する理想出力時刻 $F_{i,1}$ を合流制御部8へバイパスするよう選択部63を制御し、デキュー番号 i が自担当のキュー番号 i と一致しなければ、次に後計算部5（F計算部51-i）から到着する理想出力時刻 $F_{i,1}$ をバイパスせずに最小値選択部7へ出力するよう選択部8を制御することが実現されている。

10 【0072】従って、バイパス制御部6全体でみると、バイパス制御部6は、デキュー番号 i と対応するキュー# i の理想出力時刻 $F_{i,1}$ を廃棄するとともに、それ以外のキュー# k （ $k \neq i$ ）の理想出力時刻 $F_{k,1}$ を最小値選択部7に出力し、且つ、デキュー番号 i 対応のバイパス状態（ S_i ）をON状態に設定し、デキュー番号 i 以外のキュー# k に対応するバイパス状態（ S_k ）をOFF状態に設定していることになる。

20 【0073】次に、最小値選択部7は、上記のバイパス制御部6（バイパス制御回路61-i）から入力されてくる理想出力時刻 F_i （ただし、合流制御部8へバイパスされたものは除く）の中から、最も値の小さい理想出力時刻 F_i のキュー番号 i とその値 F_i を出力するもので、例えば図2に示すように、それぞれ値の小さい方を選択出力する（ $n-1$ ）個の比較器71により、トーナメント方式で各理想出力時刻 F_i を比較することで、最も値の小さい理想出力時刻 F_i を出力するようになって

30 【0074】そして、合流制御部8は、上記のバイパス制御部6によってバイパスされてくる理想出力時刻 F_i と、上記の最小値選択部7で選択された理想出力時刻 F_i とを大小比較することにより、値の小さい方の理想出力時刻 F_i を選択するもので、選択された理想出力時刻 F_i のキュー番号 i がデキュー指示としてキュー管理部2（デキュー処理部23）、中間結果記憶部4、バイパス制御部6にフィードバックされるとともに、その値 F_i が直前に出力した先頭パケットの出力時刻 $F_{i,0}$ （次の計算のための入力値）として中間結果記憶部41（中間データ記憶部41-i）にフィードバックされて記憶されるようになっている。

40 【0075】つまり、本実施形態のパケット・スケジューラ1は、或るキュー# i の先頭パケットがデキューされると、パイプライン処理により、それまでに中間結果記憶部4から後計算部5に既に読み出されて後計算部5で計算済みの2番目のパケットについての理想出力時刻 F_i をバイパス制御部6によって合流制御部8へバイパスすることで、その理想出力時刻 F_i と、デキューされなかった他のキュー# k の先頭パケットについての理想出力時刻 F_k とが同じタイミングで合流制御部8に入力されて大小比較されるようにしているのである。

50 【0076】これにより、合流制御部8では、同じキュー

一# i から連続して先頭バケットがデキューされる場合にも、デキューされなかったキュー# k の先頭バケットとデキューされたキュー# i の先頭バケットの次のバケットとの理想出力時刻比較を必ずパイプライン段数1段分の処理時間($p\tau$)毎に行なうことができる。即ち、同じキュー# i から $p\tau$ 間隔で連続して先頭バケットをデキューすることが可能となるのである。

【0077】以下、上述のごとく構成された本実施形態のバケット・スケジューラ1の動作について詳述する。ただし、以下の説明では、説明の簡単化のために、例えば図5に示すように、キュー# i の数(n)が“2”の場合を例にする。なお、この図5において、“Q1”はキュー# 1、“Q2”はキュー# 2、“P1”は先頭バケットの情報(中間データ)を表わす。即ち、“Q1, P1”はキュー# 1 の先頭バケットの情報を意味し、“Q2, P1”はキュー# 2 の先頭バケットの情報を意味する。これらの記号の意味は、以降で使用する図6～図16においても同様である。

【0078】(1) 初期状態〔バケット出力(デキュー指示)が無い状態〕からデキュー指示(バケット出力)が発生するまでの動作説明(図6～図8)

まず、図6に示すように、初期状態では、バケットの出力がないので、デキュー指示も発生していない。このとき、バイパス制御部6での各キュー# 1, # 2に対応するバイパス制御回路61-1, 61-2のバイパス状態(状態変数S)は全てOFF状態にリセットされている。

【0079】かかる状態で、キュー管理部2の各キュー# 1, # 2にそれぞれ送信待ちのバケットが存在しているとすると、システムクロック〔一定の時間間隔($p\tau$)〕に従って、キュー管理部2のエンキュー処理部21から各キュー# 1, # 2の先頭バケットと2番目のバケットの情報が前計算部3に出力され、前計算部3にて、各キュー# 1, # 2のそれぞれについて、先頭バケットと2番目のバケットに対する β の値が計算されるとともに、先頭バケットについての γ の値が計算される。

【0080】そして、得られた2バケット分の β の値は、それぞれ、キュー# 1, # 2毎に、中間結果記憶部4の対応する中間データ記憶部41-1, 41-2に、先頭バケット(P1)及び2番目のバケット(P2)についての中間データの一部として一旦保持される。中間結果記憶部4は、各中間データ記憶部41-1, 41-2に保持されている2バケット分の中間データのうち先頭バケットについての中間データの組をそれぞれ後段の後計算部5へ送る(符号S1参照)。なお、バケットの無い空き状態のキュー# i については、バケットが存在しない識別情報を送って、以降の段では処理しない。

【0081】その後、後計算部5(F計算部51-1, 51-2)での計算を経て、各キュー# 1, # 2の先頭バケットについて求められた理想出力時刻 F_1, F_2 の組

がバイパス制御部6に到着したとする。ここで、この場合、バイパス制御部6(バイパス制御回路61-1, 61-2)における状態変数S i はそれぞれOFF状態であるので、バイパス制御部6は、いずれも、到着した理想出力時刻 F_1, F_2 をバイパスしないように制御する(図4に示す状態遷移65-4参照)。

【0082】即ち、図7に示すように、一定の時間間隔($p\tau$)をトリガとして、バイパス制御部6は、状態変数S i = OFF状態のキュー# 1, # 2に対して、到着した理想出力時刻 F_1, F_2 をそれぞれ最小値選択部7に出力する(符号S2参照)。一方、このとき、中間結果記憶部4では、デキュー指示が発生していないので、前回送ったものと同じ先頭バケットについての中間データの組をそれぞれ後計算部5へ送る(符号S3参照)。

【0083】そして、最小値選択部7では、入力してきた理想出力時刻 F_1, F_2 のうち、値の小さい方の理想出力時刻 F_1, F_2 (仮に、 F_1 とする)を選択して、合流制御部8へ出力する。この処理と平行して、後計算部5からは、図8に示すように、次の中間データの組が出力されて、バイパス制御部6に入力される。このとき、何れのバイパス制御回路61-1, 61-2の状態変数SもOFF状態であるので、到着した中間データの組は、それぞれ、対応する一時記憶部62に記憶される(符号S4, 図4に示す状態遷移65-4参照)。

【0084】一方、このとき、合流制御部8では、最小値選択部7の出力とバイパス制御部6の出力とを合流して値の小さい方を選択するが、この場合は、バイパス制御部6からバイパスされてくるデータが無いので、最小値選択部7から入力された理想出力時刻 F_1 をそのまま選択する(符号S5参照)。そして、デキュー指示(De QueueNo=1)を前段(キュー管理部2, 中間結果記憶部3, バイパス制御部6)にフィードバックするとともに、選択した理想出力時刻 F_1 を次の計算のために中間結果記憶部4へフィードバックする。

【0085】(2) 同じキュー# i から連続してバケットが出力される場合の動作説明

次に、以下では、同じキュー# i に対して連続してデキュー指示が発生して、同じキュー# i から2つのバケットが連続して出力される場合の動作について説明する。

まず、図9に示すように、図8により上述のごとく、合流制御部8においてキュー# 1の先頭バケットについての理想出力時刻 F_1 が選択されて、デキュー指示(De QueueNo=1)が、キュー管理部2, 中間結果記憶部4及びバイパス制御部6にフィードバックされたとする(キュー# 1の先頭バケットが出力リンクに出力されたとする)。

【0086】すると、図10に示すように、キュー管理部2では、キュー# 1から次(3番目)のバケット(P3)の情報をキュー管理テーブル221から読み出して前計算部3に出力する(符号S6参照)。また、中間結

果記憶部 4 では、キュー # 1 の 2 番目のバケットについての中間データ (Q 1, P 2) を新たな先頭バケットについての中間データ (Q 1, P 1) として読み出して、デキュー指示 (DeQueueNo=1) を受けたキュー # 1 以外のキュー # 2 の先頭バケットについての中間データ (Q 2, P 1) と組で、後計算部 5 に出力する (符号 S 7 参照)。このとき、キュー # 1 の 2 番目のバケットについての中間データ (Q 1, P 2) は、先頭バケットについての中間データ (Q 1, P 1) として一時記憶部 6 2 に記憶される (符号 S 7' 参照)。

【0087】さらに、バイパス制御部 6 では、上記のデキュー指示 (DeQueueNo=1) により、キュー # 1 に対応する一時記憶部 6 2 に記憶されている理想出力時刻 F_1 を破棄して (無効にして; 図 9 参照)、キュー # 1 対応の状態変数 S を ON 状態に設定する (図 4 に示す状態遷移 6 5-2 参照) とともに、キュー # 1 以外の理想出力時刻 F_2 を最小値選択部 7 に出力する (符号 S 8 参照)。このとき、キュー # 1 以外の状態変数 S は OFF 状態のままである (図 4 に示す状態遷移 6 5-3 参照)。

【0088】その後、図 11 に示すように、上記の符号 S 8 で示すキュー # 2 の先頭バケットについての理想出力時刻 F_2 は、最小値選択部 7 に入力され、最小値選択が行なわれる (符号 S 9 参照)。一方、このとき、後計算部 5 での処理遅延時間を経て、図 10 において中間結果記憶部 4 から読み出された中間データ (符号 S 7 参照) に基づいて求められた理想出力時刻 F_1 , F_2 がバイパス制御部 6 の入力に到着する (符号 S 10 参照)。

【0089】そして、図 12 に示すように、バイパス制御部 6 では、キュー # 1 対応の状態変数 S = ON 状態であるため、到着した理想出力時刻 F_1 を一時記憶部 6 2 に記憶する (符号 S 11 参照) とともに、合流制御部 8 へバイパスする (符号 S 12, 図 4 に示す状態遷移 6 5-7 参照)。なお、このとき、状態変数 S = OFF 状態のキュー # 2 の先頭バケットについての理想出力時刻 F_2 は、バイパスせずに一時記憶部 6 2 に記憶する (符号 S 13, 図 4 に示す状態遷移 6 5-4 参照)。

【0090】合流制御部 8 では、最小値選択部 7 から送られてきたキュー # 1 以外のキュー # 2 に対応する理想出力時刻 F_2 と、バイパス制御部 6 からバイパスされてくるキュー # 1 の先頭バケット (元々は 2 番目のバケット) についての理想出力時刻 F_1 の中から最小のものを選んで出力する。仮に、キュー # 1 の先頭バケットについての理想出力時刻 F_1 が選ばれたら、つまり、同じキュー # 1 に対するデキュー指示 (DeQueueNo=1) が連続して発生した場合、そのデキュー指示 (DeQueueNo=1) により、以下の処理が行なわれる。

【0091】即ち、図 13 に示すように、まず、キュー管理部 2 では、キュー管理テーブル 2 2 1 からキュー # 1 の次 (3 番目) のバケットの情報を読み出して、前計算

部 3 に出力する (符号 S 15 参照)。一方、中間結果記憶部 4 では、上記のデキュー指示 (DeQueueNo=1) により、中間データ記憶部 4 1-1 からキュー # 1 の 2 番目のバケットに関する中間データ (Q 1, P 2) を新たな先頭バケットについての中間データ (Q 1, P 1) として、他のキュー # 2 の先頭バケットに関する中間データ (Q 2, P 1) との組で後計算部 5 に出力する (符号 S 16 参照) とともに、上記の中間データ (Q 1, P 2) を新たな先頭バケットについての中間データ (Q 1, P 1) として記憶する (符号 S 17 参照)。

【0092】また、バイパス制御部 6 では、上記のデキュー指示 (DeQueueNo=1) により、キュー # 1 に対応する一時記憶部 6 2 の (理想出力時刻 F_1) を無効にする (符号 S 18 参照)。なお、このとき、キュー # 1 対応の状態変数 S は ON 状態のままである (図 4 に示す状態遷移 6 5-5 参照)。一方、他のキュー # 2 の理想出力時刻 F_2 は最小値選択部 7 に出力される (符号 S 19 参照)。このとき、キュー # 2 対応の状態変数 S は、デキュー番号が一致しないので、OFF 状態のままである (図 4 に示す状態遷移 6 5-3 参照)。

【0093】以降は、合流制御部 8 にて他のキュー # 2 の先頭バケットについての理想出力時刻 F_2 が選択されてデキュー指示 (DeQueueNo=2) が発生するまで、図 11~図 13 により上述したパイプライン処理が繰り返し実施されて、同じキュー # 1 から連続してバケットがデキューされる。

(3) 前回と異なるキュー # i に対するデキュー指示が発生した場合の動作説明

さて、以下では、図 12 に示す状態において、合流制御部 8 にて、キュー # 1 の先頭バケットについての理想出力時刻 F_1 よりもキュー # 2 の先頭バケットについての理想出力時刻 F_2 の方が小さく、キュー # 2 に対するデキュー指示 (DeQueueNo=2) が発生した場合 (符号 S 14 参照) の動作を説明する。

【0094】即ち、図 14 に示すように、合流制御部 8 にて、キュー # 2 の先頭バケットについての理想出力時刻 F_2 が選ばれて出力されたとする (符号 S 19' 参照) と、キュー # 2 に対するデキュー指示 (DeQueueNo=2) が、キュー管理部 2, 中間結果記憶部 4, バイパス制御部 6 にそれぞれフィードバックされる。すると、図 15 に示すように、キュー管理部 2 では、上記のデキュー指示 (DeQueueNo=2) により、キュー # 2 から次 (3 番目) のバケットの情報を読み出して前計算部 3 に出力する (符号 S 20 参照)。一方、中間結果記憶部 4 では、キュー # 2 の 2 番目のバケットについての中間データをキュー # 2 の新たな先頭バケットについての中間データとして読み出して、他のキュー # 1 の先頭バケットについての中間データと組で、後計算部 5 に出力する (符号 S 21 参照) とともに、キュー # 2 の先頭バケットについての中間データを先頭バケットについての中間データ

10

20

30

40

50

として記憶する（符号 S 2 2 参照）。

【0095】また、このとき、バイパス制御部 6 では、上記のデキュー指示（DeQueueNo=2）により、キュー # 2 に対応するバイパス制御回路 6 1-2 の一時記憶部 6 2 に保持されているデータ（理想出力時刻 F_2 ）を破棄するとともに、キュー # 2 に対応する状態変数 S を ON 状態に設定する（図 4 に示す状態遷移 6 5-2 参照）。なお、他のキュー # 1 に対応するバイパス制御回路 6 1-1 では、デキュー指示（DeQueueNo=2）によるデキュー番号が一致しないので、状態変数 S は OFF 状態のまま、一時記憶部 6 2 に保持中のデータ（理想出力時刻 F_1 ）を最小値選択部 7 に出力する（符号 S 2 3 及び図 4 に示す状態遷移 6 5-5 参照）。

【0096】最小値選択部 7 では、この場合、バイパス制御部 6 からの入力がないので、図 1 6 に示すように、キュー # 1 に対応する理想出力時刻 F_1 を合流制御部 8 へ出力する（符号 S 2 4 参照）。一方、このとき、後計算部 5 での処理遅延時間を経て、中間結果記憶部 5 から出力されたキュー # 1、# 2 の次のバケットについての理想出力時刻 F_1 、 F_2 の組が、バイパス制御部 6 の入力

に到着する。
【0097】バイパス制御部 6 では、キュー # 2 の状態変数 S = ON 状態であるため、到着した理想出力時刻 F_2 を合流制御部 8 へバイパスさせる（符号 S 2 5 参照）。ただし、このときバイパスした理想出力時刻 F_2 は一時記憶部 6 2 に保持しておく（図 4 に示す状態遷移 6 5-7 参照）。また、キュー # 2 以外の状態変数 S = OFF 状態のキュー # 1 に対応する理想出力時刻 F_1 は、バイパスされずに対応する一時記憶部 6 2 に保持される（図 4 に示す状態遷移 6 5-4 参照）。

【0098】そして、合流制御部 8 では、最小値選択部 6 からのデータ（理想出力時刻 F_1 ）とバイパス制御部 6 からバイパスされてきたデータ（理想出力時刻 F_2 ）の中から最小のものを選んで出力する。なお、仮に、キュー # 1 に対応する理想出力時刻 F_1 が選ばれたら、次は、図 9 からの処理が行なわれ、キュー # 2 に対応する理想出力時刻 F_2 が再度選ばれたら、図 1 4 からの処理が行なわれる。

【0099】ここで、最初にキュー # 1 に対するデキュー指示（DeQueueNo=2）が発生した後、次に、キュー # 2 に対するデキュー指示（DeQueueNo=1）が発生する場合の動作をタイムチャートにして表わすと図 1 7 のようになる。なお、この図 1 7 において、P 1 はキュー # i の先頭バケットを表わし、(Q_{i,n}) は連続的にみたときのキュー # i の n 番目のバケットについてのデータを表わす。

【0100】即ち、最小値選択部 7 での比較の結果、キュー # 1 に対応する理想出力時刻 F_1 が選ばれて、デキュー指示（DeQueueNo=1）が発生したとする（矢印 t 1、t 2 参照）。なお、このとき、バイパス制御部 6 に

は、後計算部 5 での処理遅延時間を経て、次の先頭バケットについての中間データの組が既に保持されている（矢印 t 3 参照）。

【0101】上記のデキュー指示（DeQueueNo=1）は、バイパス制御部 6（バイパス制御回路 6 1-1、6 1-2）、中間結果記憶部 4（中間データ記憶部 4 1-1、4 1-2）、キュー管理部 2 のそれぞれにフィードバックされ（矢印 t 4 参照）、まず、バイパス制御部 6 では、キュー # 1 に対応する理想出力時刻 F_1 を無効にする（出力しないことで次に到着するデータで上書きさせる）一方、キュー # 1 以外の理想出力時刻 F_2 を最小値選択部 7 へ出力する（矢印 t 5 参照）。

【0102】また、中間結果記憶部 4 では、上記のデキュー指示（DeQueueNo=1）により、キュー # 1 については 2 番目のバケット（P 2）の中間データを新たな先頭バケットについての中間データとして後計算部 5（F 計算部 5 1-1）に出力する（矢印 t 6 参照）一方、その他のキュー # 2 については再度同じ先頭バケット（P 1）についての中間データを後計算部 5（F 計算部 5 1-2）に出力する（矢印 t 7 参照）。

【0103】さらに、キュー管理部 2 では、上記のデキュー指示（DeQueueNo=1）により、キュー # 1 の 3 番目のバケット（Q 1、P 3）の情報がキュー管理テーブル 2 2 1 から読み出されて前計算部 3 へ出力している（矢印 t 8 参照）。その後、合流制御部 8 には、最小値選択部 7 からのデキュー指示（DeQueueNo=1）の無かったキュー # 2 に対応する理想出力時刻 F_2 が到着する（矢印 t 9 参照）とともに、後計算部 5（F 計算部 5 1-1）で求められたキュー # 1 の 2 番目のバケットについての理想出力時刻 F_1 が新たな先頭バケットについての理想出力時刻 F_1 としてバイパス制御部 6 でバイパスされて到着する（矢印 t 10、t 11 参照）。

【0104】このとき、中間結果記憶部 4（中間データ記憶部 4 1-1）には、キュー # 1 の 3 番目のバケットについての前計算部 3 による計算結果が到着している（矢印 t 8' 参照）。また、バイパスされなかったキュー # 2 の先頭バケットの理想出力時刻 F_2 はバイパス制御部 6 で一時的に保持される（矢印 t 10' 参照）。そして、合流制御部 8 は、到着した各理想出力時刻 F_1 、 F_2 の中で最小のものを選んで出力する。仮に、今度は理想出力時刻 F_2 の方が小さかったとすると、デキュー指示（DeQueueNo=2）が発生し（矢印 t 12 参照）、これが再び、バイパス制御部 6（バイパス制御回路 6 1-1、6 1-2）、中間結果記憶部 4（中間データ記憶部 4 1-1、4 1-2）、キュー管理部 2 のそれぞれにフィードバックされる（矢印 t 13 参照）。

【0105】すると、まず、バイパス制御部 6 では、今度は、キュー # 2 に対応する理想出力時刻 F_2 を無効にする（出力しないことで次に到着するデータで上書きさせる）一方、キュー # 2 以外の理想出力時刻 F_1 を最小

値選択部 7 へ出力する（矢印 t 1 4 参照）。一方、中間結果記憶部 4 では、上記のデキュー指示（DeQueueNo=2）により、今度は、キュー # 2 については 2 番目のバケット（P 2）の中間データを新たな先頭バケットについての中間データとして後計算部 5（F 計算部 5 1-2）に出力する（矢印 t 1 5 参照）一方、その他のキュー # 1 については再度同じ先頭バケット（P 1）についての中間データを後計算部 5（F 計算部 5 1-1）に出力する（矢印 t 1 6 参照）。

【0106】また、キュー管理部 2 では、上記のデキュー指示（DeQueueNo=2）により、キュー # 2 の 3 番目のバケット（Q 2, P 3）の情報がキュー管理テーブル 2 2 1 から読み出されて前計算部 3 へ出力している（矢印 t 1 7 参照）。その後、合流制御部 8 には、最小値選択部 7 からのデキュー指示（DeQueueNo=2）の無かったキュー # 1 に対応する理想出力時刻 F_1 が到着する（矢印 t 1 8 参照）とともに、後計算部 5（F 計算部 5 1-2）で求められたキュー # 2 の 2 番目のバケットについての理想出力時刻 F_2 が新たな先頭バケットについての理想出力時刻 F_2 としてバイパス制御部 6 でバイパスされて到着する（矢印 t 1 9, t 2 0 参照）。

【0107】このとき、中間結果記憶部 4（中間データ記憶部 4 1-2）には、キュー # 2 の 3 番目のバケットについての前計算部 3 による計算結果が到着しており（矢印 t 1 7' 参照）、バイパス制御部 6 では、バイパスされなかったキュー # 2 の先頭バケットの理想出力時刻 F_2 が一時的に保持される（矢印 t 1 9' 参照）。そして、合流制御部 8 は、到着した各理想出力時刻 F_1 、 F_2 の中で最小のものを選んでデキュー指示（DeQueueNo=i）を出力し（矢印 t 2 1 参照）、以降は、そのデキュー番号に応じて、上記と同様の処理が繰り返されてゆく。

【0108】以上のように、本実施形態のバケット・スケジューラ 1 によれば、中間結果記憶部 4 において常に先頭バケットと 2 番目のバケットの 2 つ分のバケットについての中間データが記憶されるようにしておき、デキュー指示の発生したキュー # i の 2 番目のバケットについての中間データを先頭バケットについての中間データとして後計算部 5 へ出力するので、デキューされなかったキュー # k の先頭バケットの情報とデキューされたキュー # i の 2 番目のバケットの情報を同時に、後段の処理に入力でき、キュー管理部 2 手段と前計算部 3 の処理遅延を隠蔽できる。

【0109】また、デキューされたキュー # i の 2 番目のバケットについての中間データに基づいて後計算部 5 で求められたデータ（理想出力時刻 F_i ）をバイパス制御部 6 により合流制御部 8 へバイパスすることにより、常に、各キュー # i の理想出力時刻 F_i の比較に有効なデータをパイプライン段数 1 段分の処理時間 $p\tau$ 間隔で合流制御部 8 に入力することができるので、最小値選択

部 7 での処理遅延時間を隠蔽できる。

【0110】従って、異なるキュー # i に対するデキュー指示（DeQueueNo=i）が連続して発生する場合、同じキュー # i に対するデキュー指示（DeQueueNo=i）が連続して発生する場合のいずれの場合においても、合流制御部 8 では、デキュー指示の無かったキュー # k の先頭バケットと、デキュー指示のあったキュー # i の次のバケットとについて理想出力時刻比較を必ず $p\tau$ 毎に行なうことができ、常に、約 $p\tau$ 毎にバケットのデキューを実施することができる。

【0111】この結果、デキュー指示（DeQueueNo=i）により、次の出力バケット算出（決定）までの時間を、従来の約 $3 \times p\tau$ から約 $p\tau$ の $1/3$ に短縮でき、約 3 倍のスループットの向上を実現することが可能である。また、デキュー指示（DeQueueNo=i）が発生したときには、中間結果記憶部 4 のキュー # i に対応する先頭バケットについての中間データと、バイパス制御部 6 で保持されているキュー # i に対応する理想出力時刻 F_i とをそれぞれ無効にするので、デキュー済みのバケットに対して重複してデキュー指示（DeQueueNo=i）が発生してしまうことを確実に回避することができ、本バケット・スケジューラの信頼性の向上にも大いに寄与する。

【0112】（B）その他

なお、上述したバケット・スケジューラ 1 は、図 18 に示すように、搭載先のネットワーク中継装置（バケット転送装置）11 に、複数の異なる転送レートのインタフェース（バケットインタフェース）11-1~11-6（1F1~1F6；2.5Gbps×4, 10Gbps×2）が装備されている場合、図 19 に示すように上述した各部 2~8 を 1τ で処理可能な複数のパイプライン処理ステージ（A1~A4, B1~B4, C1~C4）に分割し、図 20 に示すように、 1τ 毎のタイムスロットで、1F5→1F1→1F6→1F2→1F5→1F3→1F6→1F4 という具合に、共有メモリ 9'（図 18 参照）に保持された異なるインタフェースのバケットについてのパイプライン処理を実施するようにしてもよい。ただし、この場合、各パイプライン処理ステージ間にフリップフロップ（FF）回路などを設けてパイプライン処理の同期制御を行なう必要がある。

【0113】このようにすれば、1 つのバケット・スケジューラ 1 で、複数の異なる転送レートのインタフェース 11-1~11-6 について、上述した実施形態と同様の処理を実現することができるので、ネットワーク中継装置 11 の大幅な小型化及び性能向上に大きく寄与する。また、上述した実施形態では、中間結果記憶部 4 とバイパス制御部 6 の双方を設けているが、バイパス制御部 6 を省略した構成にしたとしても、従来（図 21 参照）に比して約 2 倍程度の処理速度（スループット）の向上を望める。さらに、上述した実施形態では、中間結果記憶部 4 に記憶しておくバケット数を“2”としてい

るが、勿論、3以上にしてもよい。

【0114】そして、本発明は、上述した実施形態に限定されるものではなく、上記以外にも本発明の趣旨を逸脱しない範囲で種々変形して実施することができる。

【0115】

【発明の効果】以上詳述したように、本発明のバケット・スケジューラによれば、少なくとも、各キューから最初に出力されるべき第1バケットデータの理想出力時刻計算に必要な第1中間データと、その次に出力されるべき第2バケットデータについての理想出力時刻計算に必要な第2中間データとを記憶しておき、最優先に送信すべきと決定された第1バケットデータと同じキューに属する第2バケットデータについての第2中間データを、そのキューの新たな第1バケットデータについての第1中間データとして他のキューに属する第1バケットデータの第1中間データとともに後計算手段へ出力するので、後計算手段での理想出力時刻計算までの処理遅延を隠蔽できて、その処理速度（スループット）を大幅に向上することができる（請求項1）。

【0116】ここで、上述のごとく最優先に送信すべきと決定された第1バケットデータについての第1中間データは無効にするのが好ましく、このようにすれば、同じ第1バケットデータが重複して最優先に送信すべきバケットデータとして決定されてしまうことを確実に回避することができるので、本バケット・スケジューラの信頼性の向上に大きく寄与する（請求項2）。

【0117】また、上述のごとく第2バケットデータについての第2中間データを新たな第1バケットデータについての第1中間データとして後計算手段へ出力した場合、上記の第2バケットデータの次に出力されるべき第3バケットデータの理想出力時刻計算に必要な第3中間データを、新たな第2バケットについての第2中間データとして記憶するようにしておけば、次の理想出力時刻の計算までに必要なデータを確保しておくことができるので、理想出力時刻計算までの処理遅延をさらに効果的に隠蔽できる（請求項3）。

【0118】さらに、後計算手段において計算された第2バケットデータの理想出力時刻情報をバイパスするバイパス制御手段をそなえることにより、それ以外の第1バケットデータの理想出力時刻情報と同時に上記の第2バケットデータの理想出力時刻情報を合流制御手段へ入力できるようにすることで、異なるキューのバケットデータが最優先に送信されるべきデータとすると連続して決定される場合、同じキューのバケットデータが最優先に送信されるべきデータであると連続して発生する場合のいずれにおいても、理想出力時刻情報の選択を一定時間間隔で行なうことができ、さらに、本バケット・スケジューラの処理速度（スループット）を大幅に向上することができる（請求項4、5）。

【0119】ここで、上記のバイパス制御手段は、最優

先に送信すべきと決定した第1バケットデータについての理想出力時刻情報を無効にするように構成するのが好ましく、このようにすれば、やはり、同じ第1バケットデータが重複して最優先に送信すべきバケットデータとして決定されてしまうことを確実に回避することができるので、本バケット・スケジューラの信頼性の向上に大きく寄与する（請求項6）。

【0120】ところで、本バケット・スケジューラは、複数のバケットインタフェースについての送信待ちのバケットデータをそれぞれ出力保証帯域の設定が可能な複数のキューにより各バケットインタフェース毎に管理する場合、上記の各手段を複数のパイプライン処理ステージに分割し、これらのパイプライン処理ステージが、それぞれ、一定時間毎に異なるバケットインタフェースについての処理を行なうように構成してもよい。このようにすれば、1つのバケット・スケジューラで、上記と同様の処理を実現することができるので、搭載先装置の大幅な小型化及び性能向上に大きく寄与する（請求項7）。

【図面の簡単な説明】

【図1】本発明の一実施形態としてのバケット・スケジューラの構成を示すブロック図である。

【図2】図1に示すバケット・スケジューラの詳細構成例を示すブロック図である。

【図3】図2に示すバイパス制御回路の構成例を示すブロック図である。

【図4】図3に示す状態制御部の動作を説明するための状態遷移図である。

【図5】図1及び図2に示すバケット・スケジューラの動作（パイプライン処理）を説明するためのブロック図である。

【図6】図1及び図2に示すバケット・スケジューラの動作（パイプライン処理；初期状態からデキュー指示が発生するまで）を説明するためのブロック図である。

【図7】図1及び図2に示すバケット・スケジューラの動作（パイプライン処理；初期状態からデキュー指示が発生するまで）を説明するためのブロック図である。

【図8】図1及び図2に示すバケット・スケジューラの動作（パイプライン処理；初期状態からデキュー指示が発生するまで）を説明するためのブロック図である。

【図9】図1及び図2に示すバケット・スケジューラの動作（パイプライン処理；同じキューに対し連続してデキュー指示が発生する場合）を説明するためのブロック図である。

【図10】図1及び図2に示すバケット・スケジューラの動作（パイプライン処理；同じキューに対し連続してデキュー指示が発生する場合）を説明するためのブロック図である。

【図11】図1及び図2に示すバケット・スケジューラの動作（パイプライン処理；同じキューに対し連続して

デキュー指示が発生する場合)を説明するためのブロック図である。

【図12】図1及び図2に示すパケット・スケジューラの動作(パイプライン処理; 同じキューに対し連続してデキュー指示が発生する場合)を説明するためのブロック図である。

【図13】図1及び図2に示すパケット・スケジューラの動作(パイプライン処理; 同じキューに対し連続してデキュー指示が発生する場合)を説明するためのブロック図である。

【図14】図1及び図2に示すパケット・スケジューラの動作(パイプライン処理; 異なるキューに対しデキュー指示が発生する場合)を説明するためのブロック図である。

【図15】図1及び図2に示すパケット・スケジューラの動作(パイプライン処理; 異なるキューに対しデキュー指示が発生する場合)を説明するためのブロック図である。

【図16】図1及び図2に示すパケット・スケジューラの動作(パイプライン処理; 異なるキューに対しデキュー指示が発生する場合)を説明するためのブロック図である。

【図17】図1及び図2に示すパケット・スケジューラの動作(パイプライン処理; 異なるキューに対しデキュー指示が発生する場合)を説明するためのタイムチャートである。

【図18】本実施形態に係るパケット・スケジューラが適用されるネットワーク中継装置の一例を示すブロック図である。

【図19】図18に示すパケット・スケジューラのパイプライン構成例を示すブロック図である。

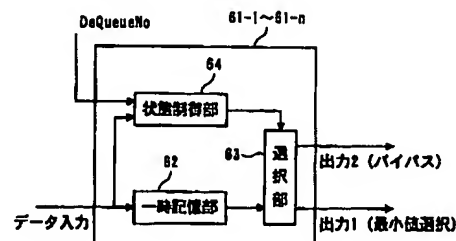
【図20】図19に示すパケット・スケジューラの動作を説明するためのタイムスロット割り当て例を示す図である。

【図21】WFQを用いたパケット・スケジューラの一例を示すブロック図である。

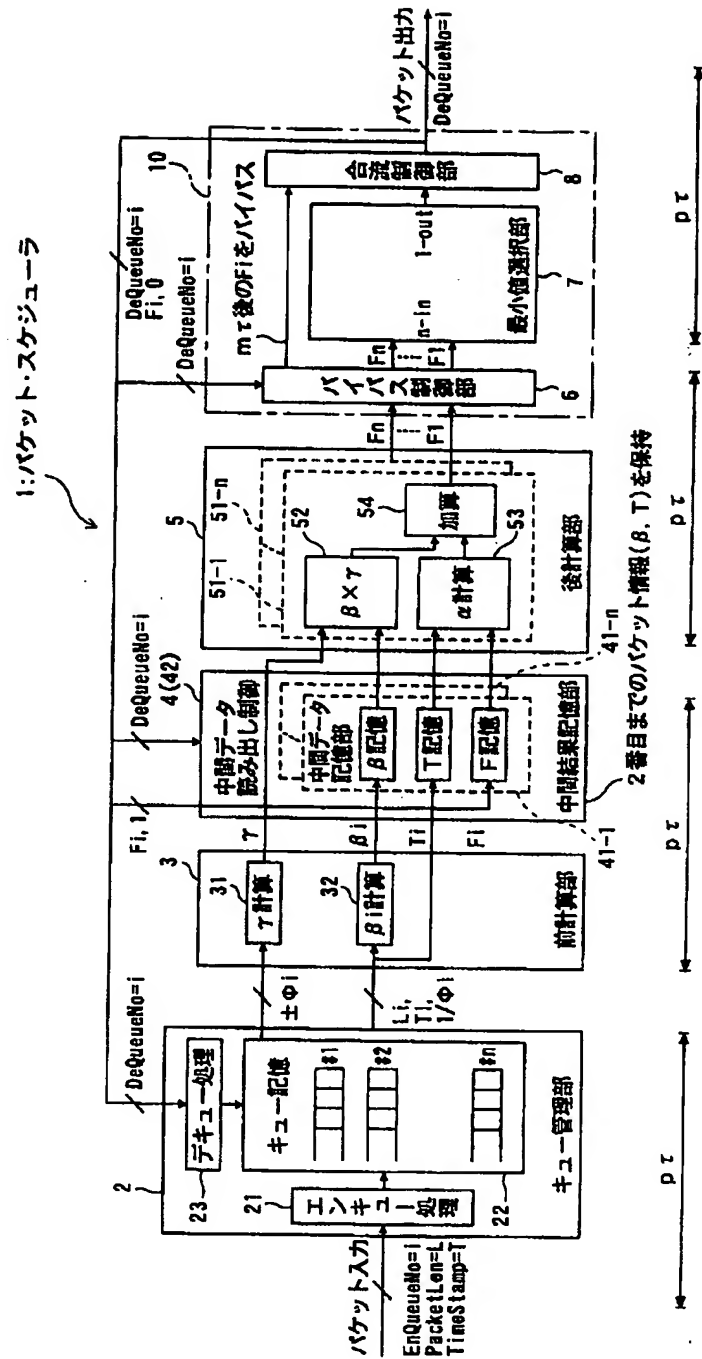
【符号の説明】

- 1 パケット・スケジューラ
- 2 キュー管理部
- 3 前計算部
- 4 中間結果記憶部
- 5 後計算部
- 6 バイパス制御部
- 7 最小値選択部
- 8 合流制御部
- 9 パケットメモリ
- 10 9' 共有メモリ
- 10 最小値決定処理部(送信パケットデータ決定手段)
- 11 ネットワーク中継装置(パケット転送装置)
- 11-1~11-6 パケットインタフェース
- 21 エンキュー処理部
- 22 キュー記憶部
- 23 デキュー処理部
- 24 エンキュー・デキュー調停部
- 25 タイマ
- 30 31 γ 計算部
- 32 β 計算部
- 41-1~41-n 中間データ記憶部
- 42 中間データ読み出し制御部
- 51-1~51-n F計算部
- 52 $\beta \times \gamma$ 計算部
- 53 α 計算部
- 54 加算部
- 61-1~61-n バイパス制御回路
- 62 一時記憶部(メモリ部)
- 63 選択部(バイパス選択手段)
- 64 状態制御部
- 71 比較器
- 221 キュー管理テーブル
- 311, 321 RAM
- 312 γ 演算器(加算器)
- 322 乗算器

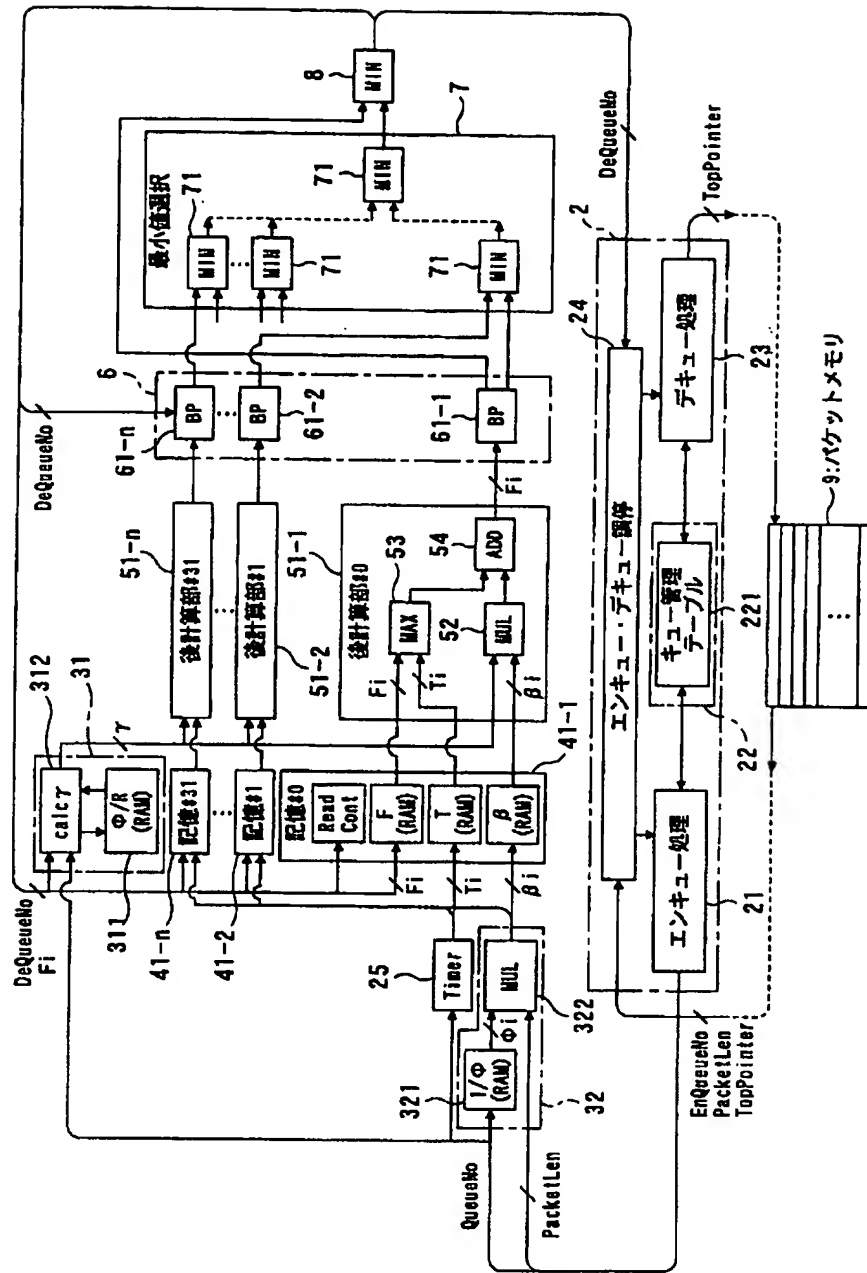
【図3】



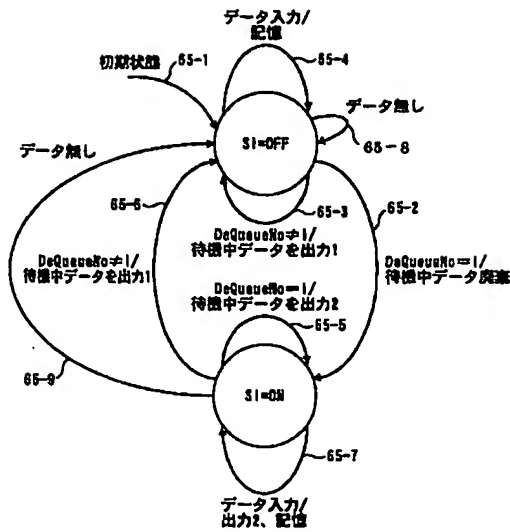
【図1】



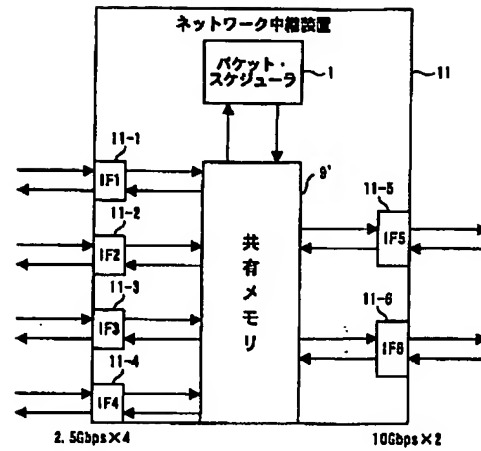
【図2】



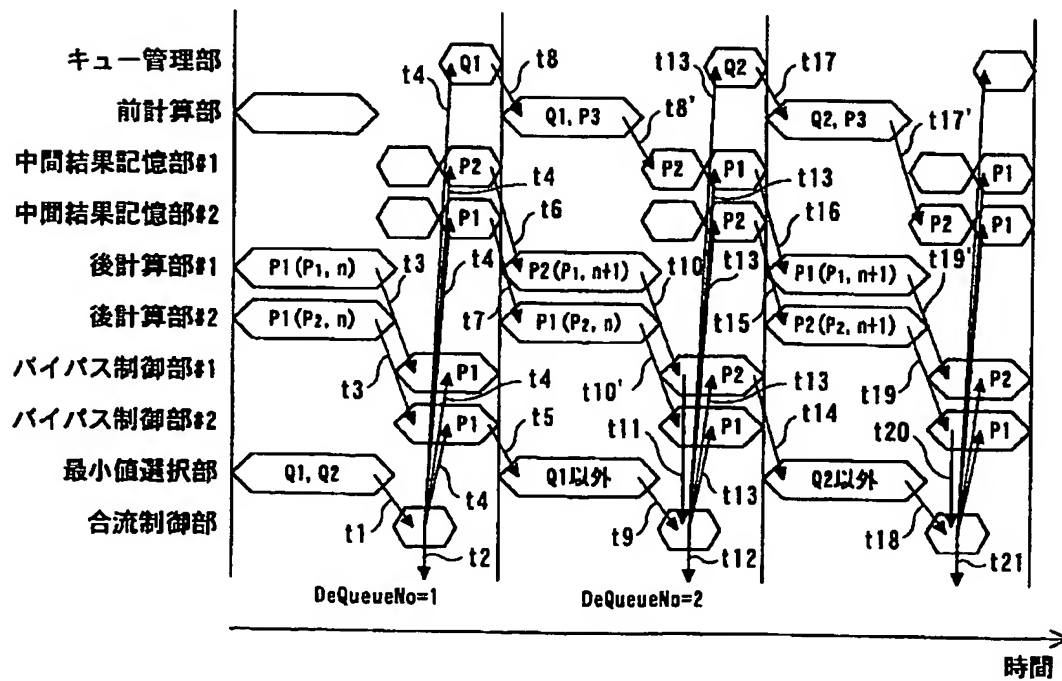
【図4】



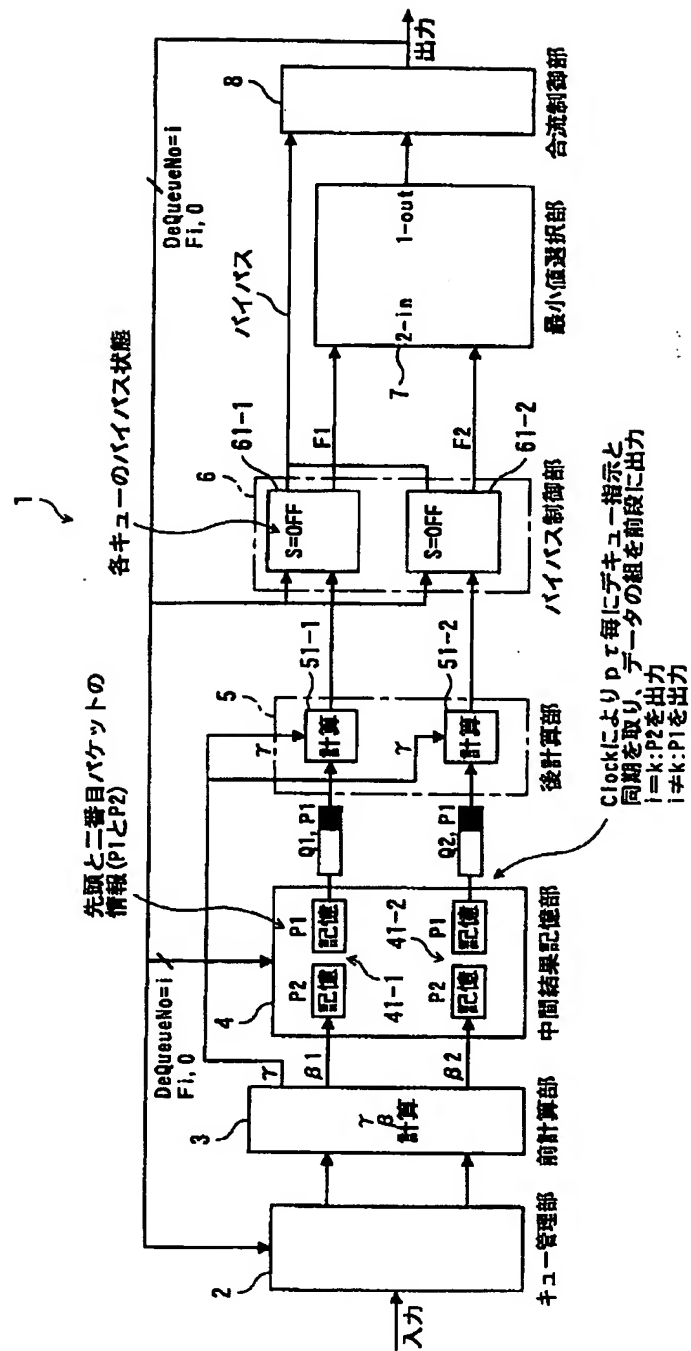
【図18】



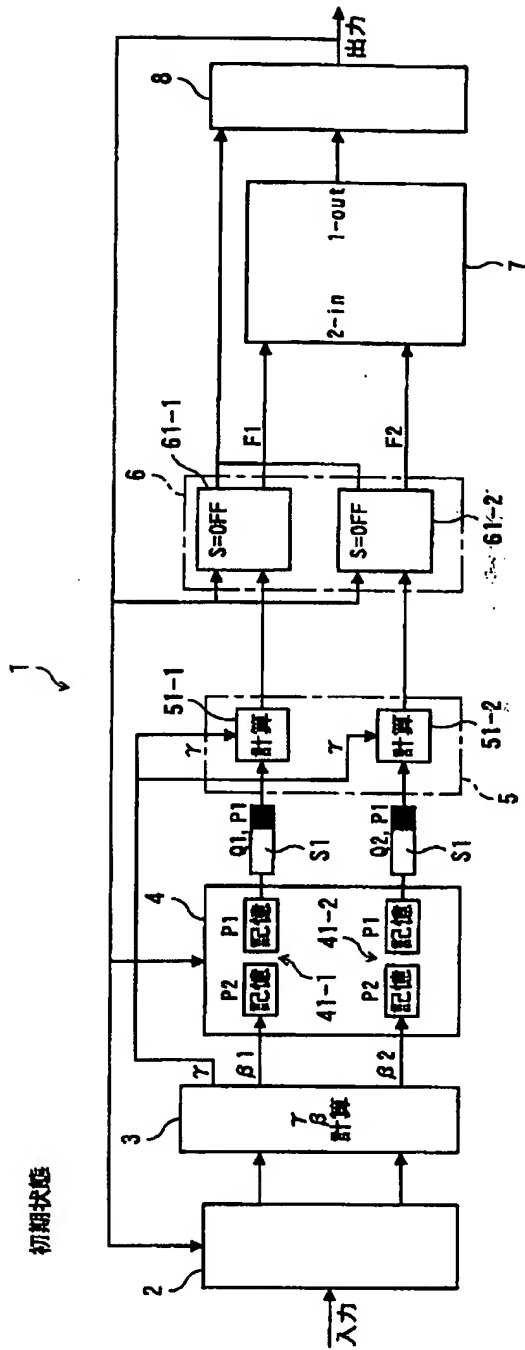
【図17】



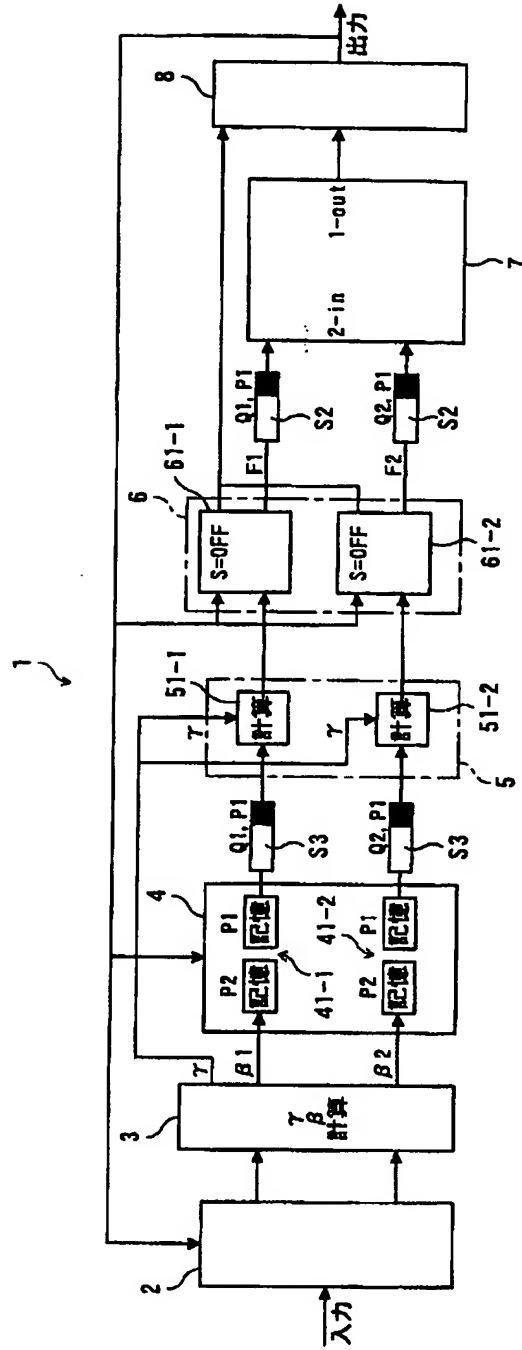
【图5】



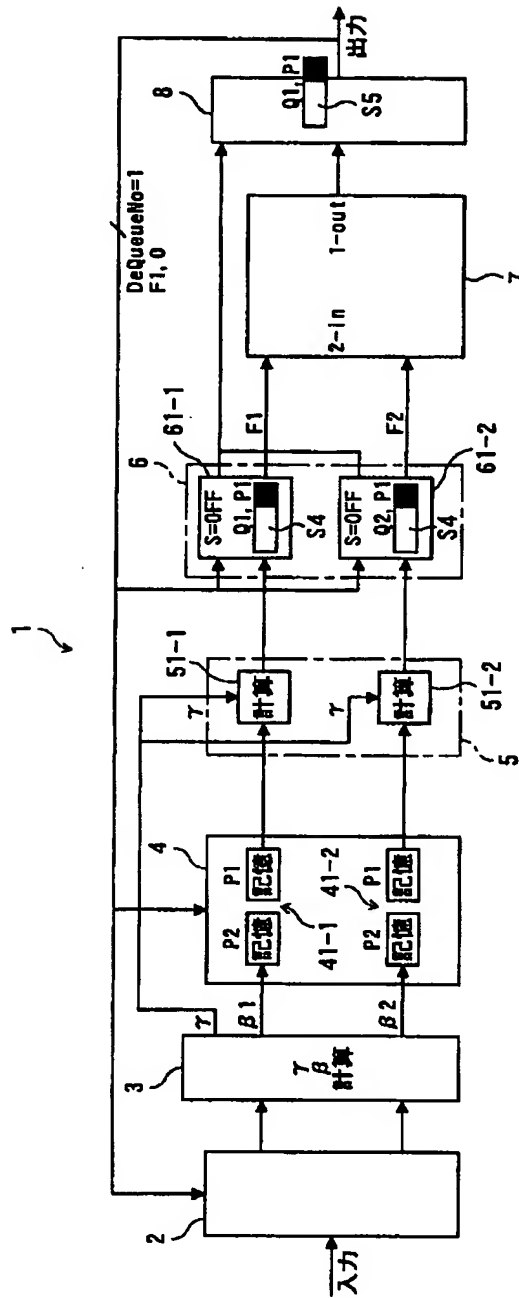
【図6】



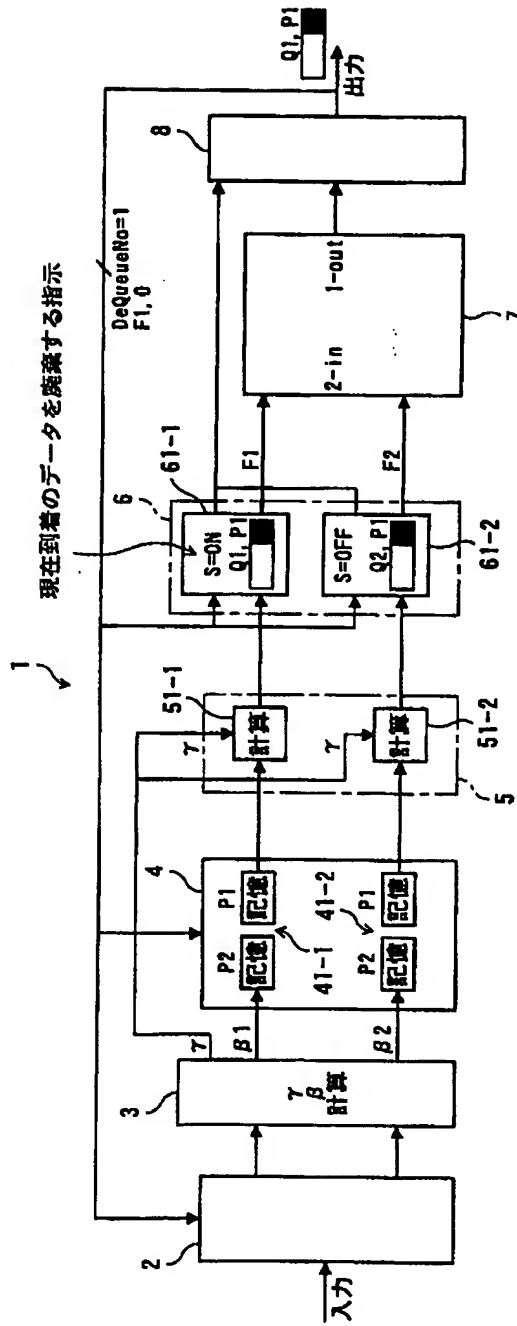
【図7】



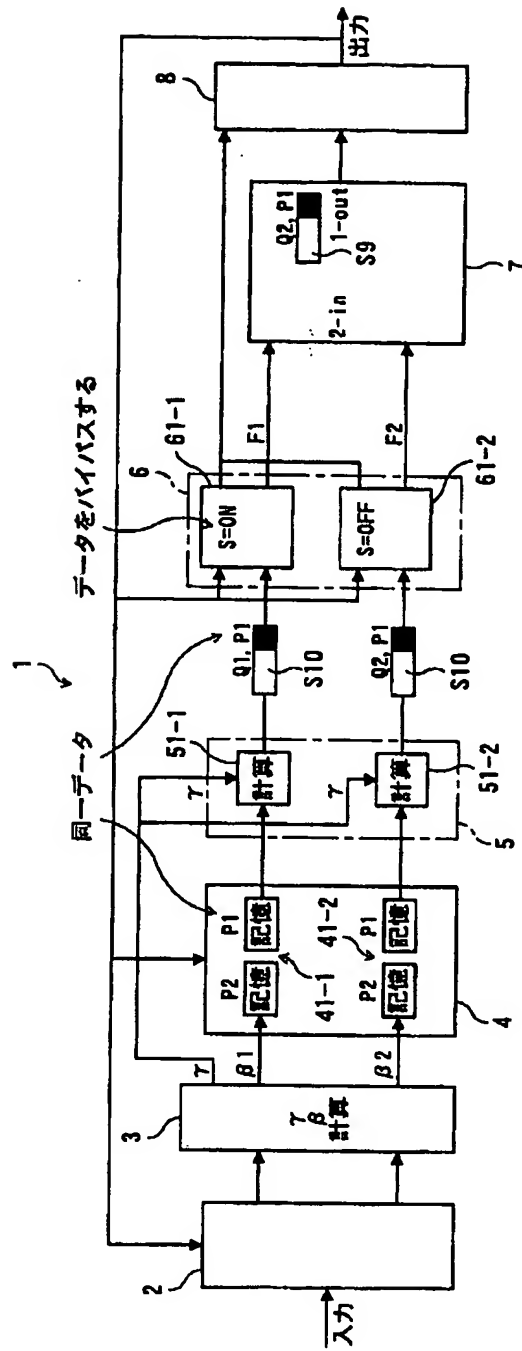
【図8】



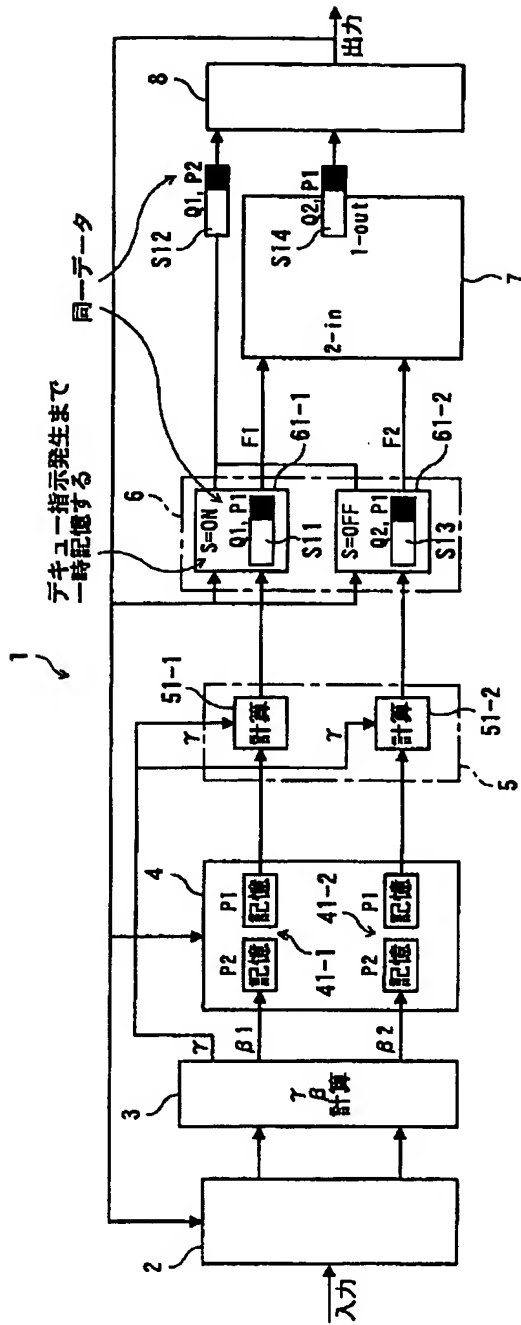
【図9】



【图 1 1】

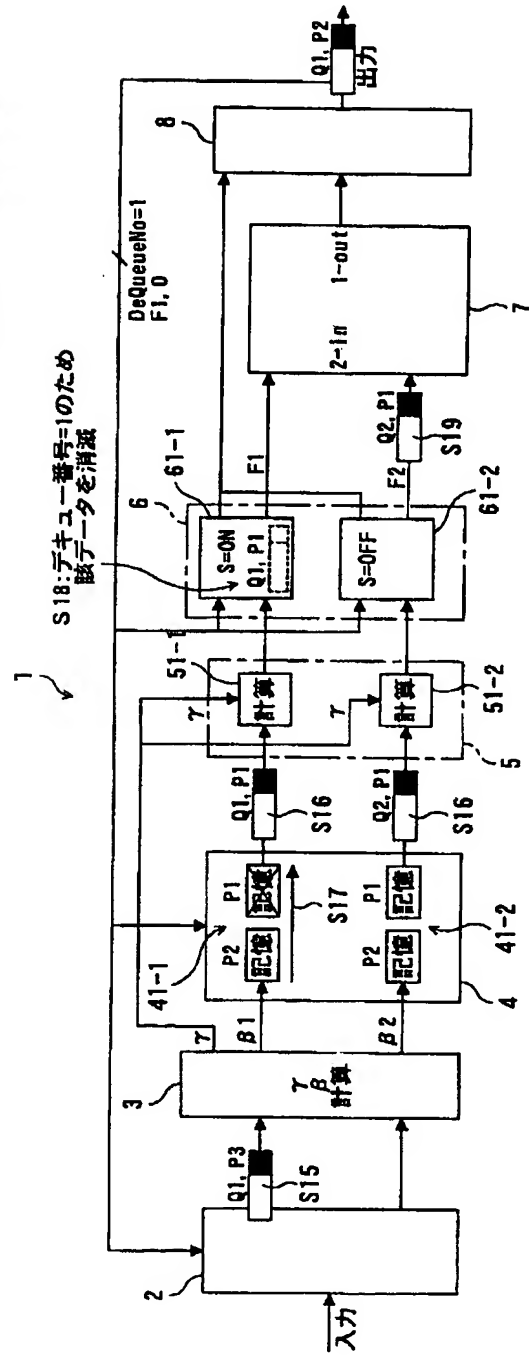


【図12】



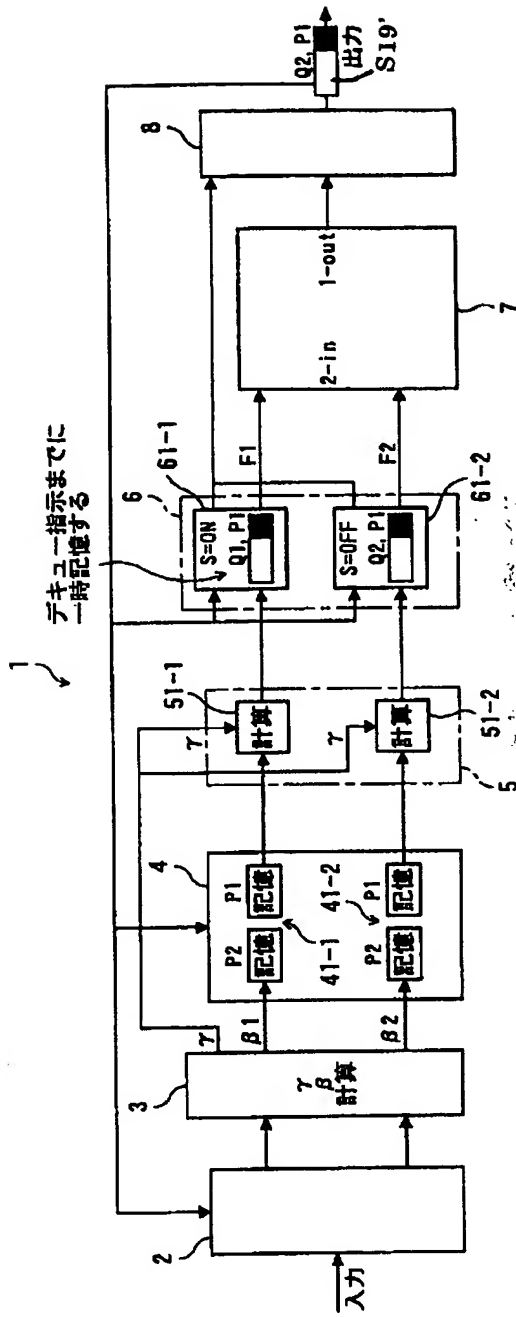
Q2から出力した場合、図14へ

【図13】

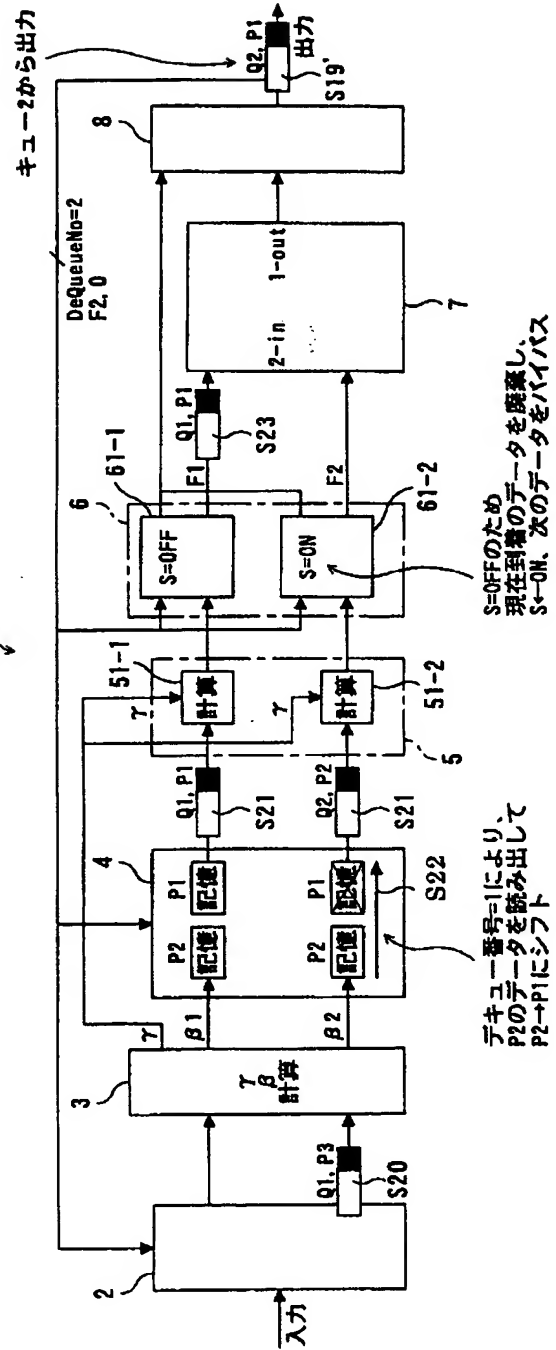


次は図10と同様

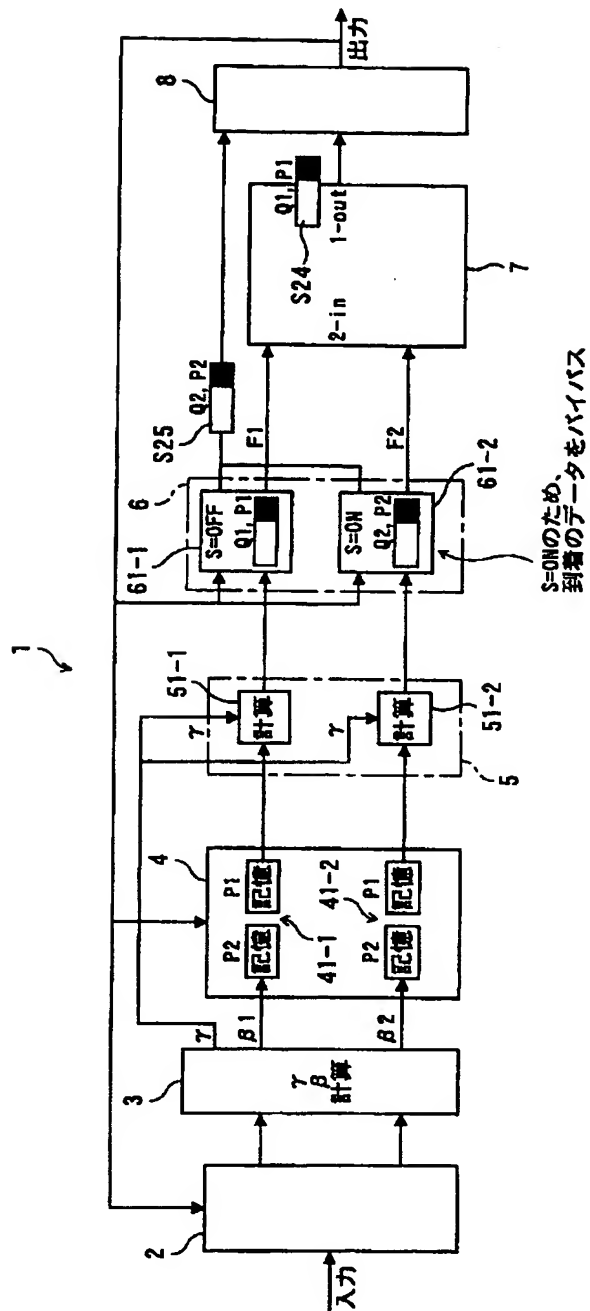
【図14】



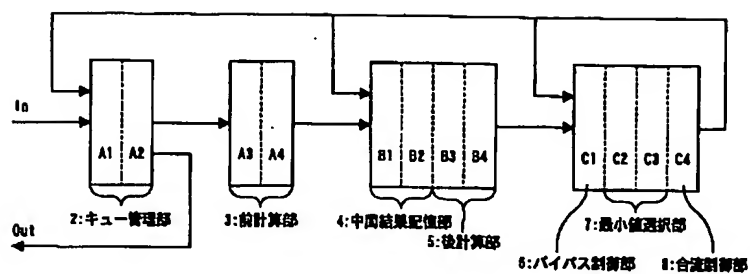
【図15】



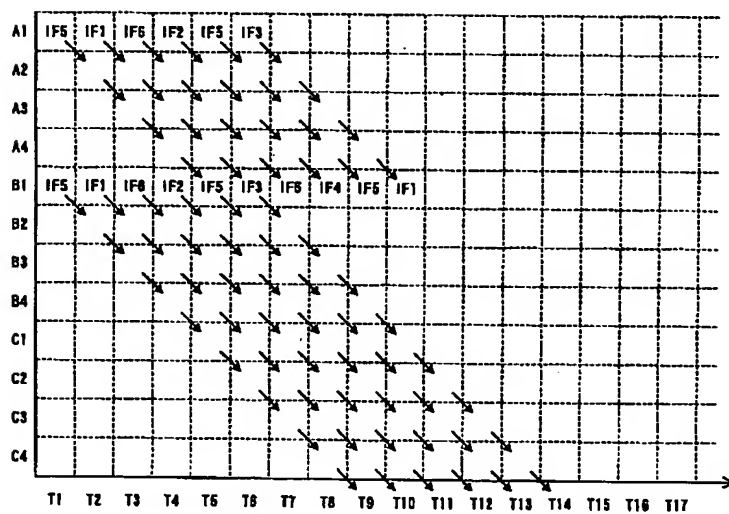
【図16】



【図19】



【図20】



最小值選擇部

蛋黃志

部
理
管
一
工
干

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.